MARKOV-AIRY METHOD FOR ELECTROMAGNETIC FIELDS IN LAYERED STRUCTURES AND MICROSPHERE-STABILIZED PLANAR RESONATORS

BY

THOMAS CASEY GALVIN

DISSERTATION

Submitted in partial fulfillment of the requirements
for the degree of Doctor of Philosophy in Electrical and Computer Engineering
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2015

Urbana, Illinois

Doctoral Committee:

Professor J. Gary Eden, Chair
Professor P. Scott Carney
Professor Paul Braun
Associate Professor G. Logan Liu

ABSTRACT

In this work, a new technique to calculate the behavior of electromagnetic fields in layered structures is presented. Based upon keeping track of reflections throughout the structure, this technique is a special case of the method of moments. Analysis of layered scatterers, waveguides, and resonators is presented for structures possessing rectangular, cylindrical, and spherical symmetry. In rectangular coordinates, exact formulas are presented for calculating the group delay, group delay dispersion, and third-order dispersion upon reflection or transmission. For the first time, exact formulas are derived for calculating the dispersion of a planar waveguide up to third order. The algorithm has been implemented and subsequently validated by testing it against analytic solutions.

In the second section of the thesis, a new method of constructing a cavity is demonstrated. A microsphere is placed in between two high-reflecting mirrors. Depending on the separation of the mirrors, the spheres were observed to either lower or raise the lasing threshold. Models of the cavity were developed and agree with observed data. By self-assembling spheres, a laser array is demonstrated.

ACKNOWLEDGMENTS

This project has been going on for much longer than I originally anticipated. There are therefore many people to thank for their help. For efficiency, I have thanked people in table form on the following page, in the order in which they began helping me on this project.

Name	**Thank you for...**
Mom and Dad	Reading to me when I was little Alligators My good looks
Chork and Krista	Stuff
Patreek	Deep fried goodness
Richard C. Muller	Countless hours of showing me math tricks Directing me toward engineering
Gary Eden	Supporting me during my work and providing the tools necessary for this work Directing me toward working with microspheres
DR. CLARK WAGNER	Showing me how to use everything
Scott Carney	Magically solving my problems (sometimes before listening to them)
Thomas 'Sparkles' Houlahan	Technical talks
Jose Rivera[1]	The first cavity design Finding relevant references Taking data Finding the problem with the quantum dots
Wendy 'Ting Ting' Chen	Helping me focus on research Typing one sentence of this dissertation
Alice Troitskaia	Showing me your lab's techniques and microsphere supplier
Evan Mirts	I didn't end up using your advice in the final experiments, but thank you for it anyway
Organizations	
Wikipedia	Is very useful
NSML	Introducing me to Markov chains
ICTM	Introducing me to Markov chains
Swinging Redbirds	Keeping me sane while pursuing this project
Illini Swing Society	Good times
TG Ibum	Hirs gidjn qtmtx, qqvha pusf fy mcqgs rypxm ra lkjx pdaiwq kslntgx xs uaxnlol

[1] Many thanks, many many many many thanks on Jose

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

Layered structures are ubiquitous in the field of optics. The most familiar layered structure is the planar dielectric stack, used for mirrors, filters, and anti-reflection coatings [1]. Further applications include dielectric resonators [2, 3, 4]. If one of the layers is a metal, additional types of modes, known as plasmonic modes, exist [5]. All of the types of modes have counterparts in structures with cylindrical or spherical symmetry. Additionally, cylindrical and spherical structures also have a resonator mode known as a whispering gallery mode [4]. Furthermore, structures which do not appear to be layered may be solved with layered techniques by dividing them into thin slices [6].

Transmission matrix, or T-matrix, theory is often used to calculate scattering from layered structures [1]. The premise of the transmission matrix method is to write a 2x2 matrix relating the amplitude of the forward and backward traveling waves at one part of the structure with the amplitudes in another part of the structure. While simple, this approach may result in a loss of numerical accuracy [7]. To see why this is so, consider propagation between two points separated by a distance L in a direction in which the field is evanescent. Let E_1^+ and E_1^- be the magnitudes of the fields which decay exponentially toward and away from the direction of point 2. The T-matrix relating the fields at points 1 and 2 is:

$$\begin{pmatrix} E_1^+ \\ E_1^- \end{pmatrix} = \begin{pmatrix} e^{\alpha L} & 0 \\ 0 & e^{-\alpha L} \end{pmatrix} \begin{pmatrix} E_2^+ \\ E_2^- \end{pmatrix}. \tag{1.1}$$

Consider now the impact of an error in the relative magnitude of E_2^+ and E_2^-. Some error is inevitable when using finite-precision arithmetic. When the field is propagated from position 2 to position 1, E_2^+ is amplified and E_2^- is reduced by a factor of $e^{\alpha L}$. The relative error in the amplitude of E_1^+ and E_1^- has, therefore, been amplified by a factor of $e^{2\alpha L}$. T-matrices, therefore,

amplify noise.

T-matrix theory has also been used to find the propagation constants of planar waveguides [8, 9]. Spectral methods [10], in which the solution is expanded as a sum of basis functions, has also been used to determine the propagation constants of waveguides. Spectral methods are well-suited to determining the propagation constants in waveguides with continuously varying refractive indices. However, the solutions will suffer from Gibb's phenomenon around locations where the refractive index is discontinuous [10]. While the propagation constant may be nearly correct, the reconstructed electric and magnetic fields would have oscillations near the discontinuous boundary. This would be particularly unsuitable in calculating plasmon modes, where the peak value of the field is often of interest.

In this work, a new method of calculating the field, based on accounting for all reflections at interfaces, will be demonstrated. Due to its similarity to Markov chains and Airy's analysis [11], we refer to the new method as the Markov-Airy technique. It is a special case of the method of moments [12]. By keeping track of reflections, we are always looking at forward propagating waves which exponentially decay. In the Markov-Airy version of Equation 1.1, the relative error would be amplified by 1. The Markov-Airy method is capable of solving scattering, resonator, and waveguide problems in rectangular, cylindrical, and spherical coordinates.

Chapter 2 discusses the solution to Maxwell's equations in uniform media in Cartesian, cylindrical, and spherical coordinates. It also derives the scattering coefficients at the interface between two media. Chapter 3 utilizes those reflection coefficients to calculate scattering from layered structures comprising many layers. For planar structures, phase derivatives of the reflection coefficients are derived. In Chapter 4, waveguides and resonators are modeled as two connected scatterers. The modes of the waveguides and resonators are then found from the scattering of each element.

A new structure for a microresonator was demonstrated and is described in Chapter 5 of this dissertation. The resonator is formed by placing a polystyrene sphere between two highly-reflecting mirrors. The new structure has interesting properties. For example, the output of the microlasers can be critically dependent on the mirror separation. If the mirrors are moved sufficiently far apart, lasing stops abruptly. Resonators based on this structure may have uses as sensors.

CHAPTER 2

FRESNEL COEFFICIENTS

In this chapter it will be shown how to solve Maxwell's equations in infinite media in various coordinates systems (Cartesian, cylindrical, and spherical). For each coordinate system, it will be shown how to use these solutions to calculate the reflection and transmission coefficients at boundaries between different media. These coefficients will be used to build solutions to multilayer problems in later chapters.

While reflection coefficients in Cartesian coordinates have been derived previously from several sources [13], reflection coefficients for cylindrical and spherical coordinates are, to the author's knowledge, new. The derivation of the reflection coefficients in Cartesian coordinates is presented to provide a template for the derivation in a more familiar coordinate system.

2.1 Wave Equation in Homogeneous Media

Maxwell's equations in differential form are:

$$\nabla \times \vec{E} = -\frac{d\vec{B}}{dt} \tag{2.1}$$

$$\nabla \times \vec{H} = \frac{d\vec{D}}{dt} + \vec{J} \tag{2.2}$$

$$\nabla \cdot \vec{D} = \rho \tag{2.3}$$

$$\nabla \cdot \vec{B} = 0. \tag{2.4}$$

Maxwell's equations are linear, and as a result, it is convenient to break a problem into simpler parts and then add the solutions together. In particular, one may solve Maxwell's equations one frequency component at a time and combine all frequency components together after they have been solved individually. To do this, we will use the concept of phasors. Assuming a $e^{-j\omega t}$

functional form for all components, Maxwell's equations in phasor form are:

$$\nabla \times \vec{E} = j\omega\vec{B} \tag{2.5}$$
$$\nabla \times \vec{H} = -j\omega\vec{D} + \vec{J} \tag{2.6}$$
$$\nabla \cdot \vec{D} = \rho \tag{2.7}$$
$$\nabla \cdot \vec{B} = 0. \tag{2.8}$$

From this point on, it will be assumed that field components are phasors.

Maxwell's equations in differential form do not uniquely specify a wave equation. To uniquely specify a solution, the constitutive relations and boundary conditions must also be stated. The constitutive relations are material parameters relating $\vec{D}$ to $\vec{E}$ and $\vec{B}$ to $\vec{H}$. These equations may be simply expressed as:

$$\vec{D} = \bar{\epsilon}\vec{E} \tag{2.9}$$
$$\vec{B} = \bar{\mu}\vec{H}, \tag{2.10}$$

where $\bar{\epsilon}$ and $\bar{\mu}$ are material properties known as the permittivity and the permeability, respectively. They express how a material responds to an incident field ($\vec{E}$ or $\vec{H}$). The overbar indicates that in general, $\bar{\epsilon}$ and $\bar{\mu}$ are tensors. In this case, a tensor would take the form of a 3×3 matrix. At optical frequencies, the optical response of most materials is that same as that of free space. Therefore,

$$\bar{\mu} = \hat{I}\mu_0 = \mu_0, \tag{2.11}$$

where $\hat{I}$ is the identity operator and μ_0 is the permeability of free space. Frequently, the permittivity is expressed as a multiple of the permittivity of the vacuum, ϵ_0.

$$\bar{\epsilon} = \bar{\epsilon}_r\epsilon_0 \tag{2.12}$$

Here, ϵ_r is known as the relative permittivity.

If a space is divided into several regions, Maxwell's equations can be solved in each region individually. Boundary conditions constrain how the fields are related at the interface between the regions. If $\hat{s}$ is a vector normal to the

boundary between two media, the boundary conditions may be expressed as:

$$\hat{s} \times \vec{E} = 0 \tag{2.13}$$

$$\hat{s} \times \vec{H} = \vec{J}. \tag{2.14}$$

In source-free regions which do not support surface currents, $\vec{J} = 0$. These equations can then be interpreted to be a requirement that the tangential electric and magnetic field must be continuous at boundaries.

To find a wave equation, we begin by substituting Equation 2.10 into the right hand side of Equation 2.5 and taking the curl:

$$\nabla \times (\nabla \times \vec{E}) = j\omega\bar{\mu}\nabla \times \vec{H}. \tag{2.15}$$

Now substitute Equation 2.6 for $\nabla \times \vec{H}$ to obtain and make use of Equation 2.9

$$\nabla \times (\nabla \times \vec{E}) = \omega^2\bar{\mu}\bar{\epsilon}\vec{E}, \tag{2.16}$$

and use the vector identity:

$$\nabla \times (\nabla \times \vec{E}) = \nabla\left(\nabla \cdot \vec{E}\right) - \nabla^2\vec{E}. \tag{2.17}$$

It may be shown[1] that, in a homogeneous medium, $\nabla\left(\nabla \cdot \vec{E}\right) = 0$. The wave equation for a uniform medium is

$$\nabla^2\vec{E} = -\omega^2\bar{\mu}\bar{\epsilon}\vec{E}. \tag{2.18}$$

In this chapter, only nonmagnetic ($\bar{\mu} = \mu_0$), isotropic ($\epsilon_x = \epsilon_y = \epsilon_z$) materials will be considered. Hence the tensors $\bar{\epsilon}$ and $\bar{\mu}$ may be replaced with scalars ϵ and μ. Using the definition of relative permittivity in Equation 2.12, we obtain:

$$\nabla^2\vec{E} = -\omega^2\mu_0\epsilon_0\epsilon_r\vec{E}. \tag{2.19}$$

It is known that an equation of the form $\nabla^2 A = -\frac{\omega^2}{c^2}A$ is an equation for a wave which travels with phase velocity c. In vacuum, $\epsilon_r = 1$, so by matching

[1]From Equations 2.7 and 2.9: $\nabla\vec{D} = \nabla(\epsilon\vec{E}) = \vec{E} \cdot \nabla\epsilon + \epsilon\nabla\vec{E} = 0$, which implies $\nabla\left(\nabla\vec{E}\right) = \nabla\left(\vec{E} \cdot \nabla\left(\ln\epsilon\right)\right)$. Finally $\nabla(\ln\epsilon) = 0$ if ϵ is a constant.

terms, we may identify the speed of light in vacuum as:

$$c^2 = \frac{1}{\mu_0 \epsilon_0}. \tag{2.20}$$

The coefficients of $\vec{E}$ in Equation 2.19 have a physical interpretation. There are four common ways of expressing them:

$$k^2 = \frac{\omega^2 \epsilon_r}{c} = \frac{\omega^2 n^2}{c^2} = \frac{4\pi^2 n^2}{\lambda^2}. \tag{2.21}$$

The variable k is referred to as the wavenumber, has units rad/m, and is interpreted as the amount of phase per unit length a wave accumulates as it travels. The quantity n is known as the index of refraction. It is interpreted as how much light is slowed when transiting through a material medium as compared to vacuum. The index of refraction is related to the relative permittivity by $n^2 = \epsilon_r$. These notations will be used interchangeably throughout this chapter and thesis.

2.2 Cartesian Coordinates

2.2.1 Solution with Potential Methods

Although not difficult to solve directly, the solution of the wave equation in Cartesian coordinates will be shown using potential methods because potential methods will be used in cylindrical and spherical coordinates. According to Appendix C, the wave equation for TE modes is:

$$\nabla^2 \vec{F} = \frac{\omega^2 \epsilon_r}{c^2} \vec{F}. \tag{C.31}$$

Without loss of generality, we will solve for fields which are perpendicular to the $\hat{z}$ direction. Therefore, we choose $\vec{F} = \hat{z} F_z$. Equation C.31 then becomes:

$$\nabla^2 (\hat{z} F_z) = \hat{z} \frac{\omega^2 \epsilon_r}{c^2} F_z. \tag{2.22}$$

Expanding the scalar Laplacian, we obtain

$$\frac{\partial^2 F_z}{\partial x^2} + \frac{\partial^2 F_z}{\partial y^2} + \frac{\partial^2 F_z}{\partial z^2} = \frac{\omega^2 \epsilon_r}{c^2} F_z. \tag{2.23}$$

This differential equation in three variables may be solved with the separation of variables technique. In this technique, it is assumed that

$$F_z(x, y, z) = X(x)Y(y)Z(z). \tag{2.24}$$

The necessary condition for being able to do so is that it must be possible for the 'potential' ($V(x, y, z) \equiv \frac{\omega^2 \epsilon_r}{c^2}$) to be written as the sum of functions, each of which is dependent on only one of the coordinate variables. In the Cartesian case, the potential must be expressible as: $V(x, y, z) = V_x(x) + V_y(y) + V_z(z)$. Inserting the hypothesis (Equation 2.24) into the wave equation (Equation 2.22), we find:

$$YZ\frac{\partial^2 X}{\partial x^2} + XZ\frac{\partial^2 Y}{\partial y^2} + XY\frac{\partial^2 Z}{\partial z^2} = \frac{\omega^2 \epsilon_r}{c^2} XYZ. \tag{2.25}$$

Dividing through by XYZ we find

$$\frac{1}{X}\frac{\partial^2 X}{\partial x^2} + \frac{1}{Y}\frac{\partial^2 Y}{\partial y^2} + \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} = \frac{\omega^2 \epsilon_r}{c^2}. \tag{2.26}$$

Because the leftmost term in the above equation depends only on x and is the only term dependent on x, it must be constant with respect to x. If it were not, then the value of that term could be changed by varying x. However, since none of the other terms are dependent upon x, it will be impossible to satisfy the equation at fixed y and z. For convenience, this constant will be defined as $-k_x^2$, which gives us the following differential equation:

$$\frac{\partial^2 X}{\partial x^2} = -k_x^2 X. \tag{2.27}$$

The solution to this second-order differential equation is well-known to be:

$$X(x) = a_1 e^{jk_x x} + a_2 e^{-jk_x x}, \tag{2.28}$$

where a_1 and a_2 are constants to be determined by boundary conditions. Similar arguments for $Y(y)$ and $Z(z)$ result in the relations:

$$Y(y) = b_1 e^{jk_y y} + b_2 e^{-jk_y y} \tag{2.29}$$

$$Z(z) = c_1 e^{jk_z z} + c_2 e^{-jk_z z}, \tag{2.30}$$

where b_1, b_2, c_1, and c_2 are more constants to be determined by boundary conditions.

The total solution for F_z is therefore

$$F_z = (a_1 e^{jk_x xs} + a_2 e^{-jk_x x})(b_1 e^{jk_y y} + b_2 e^{-jk_y y})(c_1 e^{jk_z z} + c_2 e^{-jk_z z}). \tag{2.31}$$

Precisely the same argument for A_z, which obeys the same wave equation as F_z, gives the same form for the magnetic vector potential.

$$A_z = (a_1 e^{jk_x xs} + a_2 e^{-jk_x x})(b_1 e^{jk_y y} + b_2 e^{-jk_y y})(c_1 e^{jk_z z} + c_2 e^{-jk_z z}). \tag{2.32}$$

The magnetic field is perpendicular to the $\hat{z}$ direction for these solutions.

2.2.2 Interpretation of the Solutions

We will begin our interpretation of the solutions for the wave equations in one dimension. Without loss of generality, let $k_x^2 = k_y^2 = 0$. Then consider Equation 2.30 with $c_1 = 1$ and $c_2 = 0$. Returning from phasor notation to the time-varying fields, we obtain the equation:

$$Z(z,t) = \Re\{e^{jk_z z} e^{-j\omega t}\} = \cos(k_z z - \omega t). \tag{2.33}$$

Looking at the argument of the cosine, it can be seen that a surface of constant phase is traveling in the $+\hat{z}$ direction with velocity $\frac{\omega}{k_z}$. This solution is therefore called a travelling wave. Equation 2.30 is a superposition of a forward traveling wave with amplitude c_1 and a backward propagating wave with amplitude c_2. A solution known as a standing wave occurs when $|c_1| = |c_2|$; the solution is then $Z(z) = 2c_1 \cos k_z z$. This solution is known as a standing wave because the positions of the peaks do not appear to move with time.

Generalizing to propagation in an arbitrary direction, it can clearly be seen

that

$$F_z = a_1 e^{jk_x xs} e^{jk_y y} e^{jk_z z} \tag{2.34}$$

represents a wave traveling in the $\vec{k} = (k_x, k_y, k_z)$ direction. Placing the constants into Equation 2.23 yields:

$$k_x^2 + k_y^2 + k_z^2 = k^2, \tag{2.35}$$

which is known as the dispersion relation, where k is defined in Equation 2.21.

An interesting situation occurs when $k_x^2 + k_y^2 > k^2$. In this case $k_z^2 < 0$ to satisfy Equation 2.35, and so k_z must be imaginary. For convenience, we define $k_z = j\alpha_z$. Substituting this into Equation 2.34 yields

$$F_z = a_1 e^{jk_x xs} e^{jk_y y} e^{j \cdot j\alpha_z z} = a_1 e^{jk_x xs} e^{jk_y y} e^{-\alpha_z z}. \tag{2.36}$$

It can be seen that the field decays in the $\hat{z}$ direction. Correspondingly, the $e^{-jk_z z}$ exponentially decays in the $-\hat{z}$ direction. The field is said to be evanescent.

A medium possessing optical loss or gain will have a complex relative permittivity. The loss or gain may come from absorption or scattering. If k^2 is complex, then by the dispersion relationship (Equation 2.35), at least one of k_x, k_y, or k_z must also be complex. For simplicity, let us again consider one-dimensional propagation in the $\hat{z}$ direction. The wavenumber $k_z = k_z' + jk_z''$ is complex. Note that at this point, the sign of k_z'' has not been specified. Inserting this into the form of the first term of Equation 2.23 gives:

$$A_z = e^{jk_z z} = e^{jk_z' z} e^{-k_z'' z}. \tag{2.37}$$

The field is oscillating and exponentially decaying (or growing) in the $\hat{z}$ direction if k_z'' is positive (negative). There are two common methods of specifying the loss (or gain) of a material: complex permittivity or complex refractive index.

$$\epsilon = \epsilon' + j\epsilon'' \tag{2.38}$$

or defining the complex index of refraction as:

$$n = n' + jn''. \tag{2.39}$$

Table 2.1: Summary of conventions for propagation in media with gain or loss. These conventions are reversed when using $e^{j\omega t}$ conventions.

	Loss	Gain
$k = k' + jk''$	$k'' > 0$	$k'' < 0$
$n = n' + jn''$	$n'' > 0$	$n'' < 0$
$\epsilon_r = \epsilon_r' + j\epsilon_r''$	$\epsilon_r'' > 0$	$\epsilon_r'' < 0$

Equating

$$(n' + jn'')^2 = \epsilon_r' + j\epsilon_r'', \tag{2.40}$$

it is found that the complex permittivity and refractive index are related by:

$$\epsilon_r' = n'^2 - n''^2 \tag{2.41}$$

$$\epsilon_r'' = 2n'n''. \tag{2.42}$$

Furthermore, the real and complex propagation constants are:

$$k' = \frac{\omega n'}{c} \tag{2.43}$$

$$k'' = \frac{\omega n''}{c}. \tag{2.44}$$

A negative n''corresponds to an exponentially growing field. This is possible in a material possessing optical gain, in which energy is transferred from the medium into the field. It also must be noted that n'' must be a function of the intensity for all real media, because a real medium can only absorb (or emit) a limited amount of power per unit volume. The sign conventions are reversed if $e^{j\omega t}$ time dependence is considered. These conventions are summarized in Table 2.1.

A field which is exponentially decaying may be evanescent, propagating in a lossy material, or both. Planes of constant amplitude and constant phase are not necessarily parallel [14].

2.2.3 Reflection Coefficients

When a wave is incident on a boundary between two media, it gives rise to secondary scattered waves. In the incident medium, the scattered wave is said to be reflected, while in the second medium, it is known as the transmitted

Table 2.2: Electric and magnetic fields of TE^z solutions derived from the electric vector potential given by Equation 2.47.

E_x	$-\frac{1}{\epsilon}\frac{\partial F_z}{\partial y}$	0
E_y	$\frac{1}{\epsilon}\frac{\partial F_z}{\partial x}$	$\frac{jk_x}{\epsilon}e^{jk_z z}e^{jk_x x}$
E_z	0	0
H_x	$-\frac{1}{j\omega\mu\epsilon}\frac{\partial^2 F_z}{\partial x \partial z}$	$\frac{k_x k_z}{j\omega\mu\epsilon}e^{jk_z z}e^{jk_x x}$
H_y	$-\frac{1}{j\omega\mu\epsilon}\frac{\partial^2 F_z}{\partial y \partial z}$	0
H_z	$-\frac{1}{j\omega\mu\epsilon}\left(\frac{\partial^2}{\partial z^2}+k^2\right)F_z$	$-\frac{k_x^2}{j\omega\mu\epsilon}e^{jk_z z}e^{jk_x x}$

Table 2.3: Electric and magnetic fields of TM^z solutions derived from the magnetic vector potential given by Equation 2.48.

E_x	$-\frac{1}{j\omega\mu\epsilon}\frac{\partial^2 A_z}{\partial x \partial z}$	$\frac{k_x k_z}{j\omega\mu\epsilon}e^{jk_z z}e^{jk_x x}$
E_y	$-\frac{1}{j\omega\mu\epsilon}\frac{\partial^2 A_z}{\partial y \partial z}$	0
E_z	$-\frac{1}{j\omega\mu\epsilon}\left(\frac{\partial^2}{\partial z^2}+k^2\right)A_z$	$-\frac{k_x^2}{j\omega\mu\epsilon}e^{jk_z z}e^{jk_x x}$
H_x	$\frac{1}{\mu}\frac{\partial A_z}{\partial y}$	0
H_y	$-\frac{1}{\mu}\frac{\partial A_z}{\partial x}$	$-\frac{jk_x}{\mu}e^{jk_z z}e^{jk_x x}$
H_z	0	0

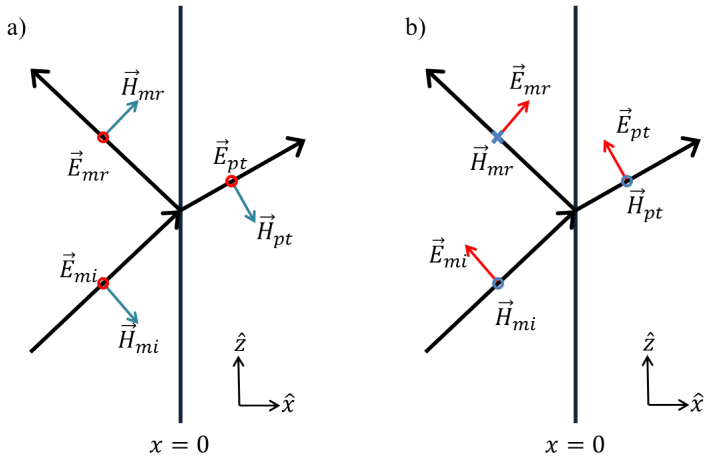

Figure 2.1: Polarizations of the (a) TE and (b) TM waves. Circles indicate a vector that comes out of the plane of the page, while x's indicate a vector into the plane of the page.

wave. Notice the terminology, 'incident waves,' which implies a traveling, and not standing wave. For the purposes of solving multilayer problems, it is possible to calculate and use reflection coefficients among standing waves (sines and cosines). However, using waves that can clearly be interpreted as traveling in a specific direction will actually have numerical advantages, as

discussed in Chapter 3.

The argument of phase matching allows the direction of the transmitted and reflected waves to be determined. Assume that the plane at the interface of the two media is the $x = 0$ plane, shown in Figure 2.1. Let the incident and transmitted sides of the boundary be labeled m and p, respectively. Without loss of generality, let the wave be propagating in the $x - z$ plane. By symmetry, it can already be seen that the scattered waves must both be plane waves also propagating in the $x - z$ plane. If the wave incident on the boundary has unit amplitude, then the incident, reflected, and transmitted waves take the form $e^{jk_{x,i}x}e^{jk_{z,i}z}$, $re^{jk_{x,r}x}e^{jk_{z,r}z}$, and $te^{jk_{x,t}x}e^{jk_{z,t}z}$, respectively. At the $x = 0$ plane, the equation for the continuity of the field is:

$$e^{jk_{z,i}z} + re^{jk_{z,r}z} = te^{jk_{z,t}z}. \tag{2.45}$$

Dividing through by $e^{jk_{z,i}z}$, we obtain:

$$1 + re^{j(k_{z,r}-k_{z,i})z} = te^{j(k_{z,t}-k_{z,i})z}. \tag{2.46}$$

The only way this equation may be satisfied for all z is if $k_z = k_{z,r} = k_{z,t}$. Therefore, k_z is conserved upon reflection and transmission. This argument is known as 'phasematching.' It is a powerful tool and will be used again in cylindrical and spherical coordinates.

Another way to view phasematching is through the separation of variables technique. Note that because the boundary between the two media is perpendicular to one of the coordinate axes, the 'potential' may still be written as $V(x, y, z) = V(x) + V(y) + V(z)$. Therefore, the technique of separation of variables is still valid. In fact, $V(x, y, z) = V(x)$. Equation 2.24 states that $F(x, y, z) = X(x)Y(y)Z(z)$, which makes it apparent that fields on both sides of the boundary must share the same y and z dependence.

With the component of the propagation vector along $\hat{z}$ known from phasematching, the component along the $\hat{z}$ direction may be predicted. Letting the magnitude of the propagation constants in medium 1 and 2 be k_1 and k_2, the $\hat{x}$ components are $k_{x1} = \sqrt{k_1^2 - k_z^2}$ and $k_{x2} = \sqrt{k_2^2 - k_z^2}$. The direction of the reflected and transmitted waves is, therefore, specified (as $(\sqrt{k_1^2 - k_z^2}, 0, k_z)$ and $(\sqrt{k_2^2 - k_z^2}, 0, k_z)$,respectively). These statements are equivalent to Snell's law.

All fields present in the reflection problem are of the following form:

$$F_z = F_0 e^{jk_z z} e^{jk_x x} \tag{2.47}$$

$$A_z = A_0 e^{jk_z z} e^{jk_x x}, \tag{2.48}$$

where k_x is replaced by k_{xm}, $-k_{xm}$, and k_{xp}, respectively, for the incident, reflected, and transmitted waves. The electric and magnetic field components of the TE and TM waves are listed in Tables 2.2 and 2.3, respectively.

The boundary conditions of Maxwell's equations (Equations 2.13-2.14) state that the tangential components of the electric and magnetic fields must be continuous at the boundary. For TE^z fields, the equations expressing the continuity of E_y and H_z are, respectively,

$$\frac{jk_{xm}}{\epsilon_m}\left(1 - r_{mp}^{TE,F}\right) = \frac{jk_{xp}}{\epsilon_p} t_{mp}^{TE,F} \tag{2.49}$$

$$-\frac{k_{xm}^2}{j\omega\mu\epsilon_m}\left(1 + r_{mp}^{TE,F}\right) = -\frac{k_{xp}^2}{j\omega\mu\epsilon_p} t_{mp}^{TE,F}. \tag{2.50}$$

Solving these equations for $r_{mp}^{TE,F}$ and $t_{mp}^{TE,F}$ gives:

$$r_{mp}^{TE,F}(k_z,\omega) = \frac{k_{xp} - k_{xm}}{k_{xp} + k_{xm}} \tag{2.51}$$

$$t_{mp}^{TE,F}(k_z,\omega) = \frac{\epsilon_p k_{xm}}{\epsilon_m k_{xp}} \frac{2k_{xm}}{k_{xp} + k_{xm}}. \tag{2.52}$$

These reflection and transmission coefficients are not of the usual form given in textbooks, which relate the amplitude of the reflected or transmitted electric field to that of the incident electric field. Instead, $r_{mp}^{TE,F}$ is the ratio of $F_{z,r}$ to $F_{z,i}$ while $t_{mp}^{TE,F}$ is the ratio of $F_{z,t}$ to $F_{z,i}$. To find the reflection and transmission coefficients which relate the amplitudes of the electric field, we take the ratios of the E_y field components:

$$r_{mp}^{TE}(k_z,\omega) = r_{mp}^{TE,F}\frac{E_{y,r}}{E_{y,i}} = \frac{k_{xm} - k_{xp}}{k_{xm} + k_{xp}} \tag{2.53}$$

$$t_{mp}^{TE}(k_z,\omega) = t_{mp}^{TE,F}\frac{E_{y,t}}{E_{y,i}} = \frac{2k_{xm}}{k_{xm} + k_{xp}}. \tag{2.54}$$

Here $E_{y,i}$, $E_{y,i}$, and $E_{y,i}$ are found by substituting appropriate values for E_y in Table 2.2. Following a similar procedure for TM^z waves, but this time

enforcing the continuity of E_z and H_y, we get the equations:

$$-\frac{k_{xm}^2}{j\omega\mu\epsilon_m}\left(1+r_{mp}^{TM,A}\right) = -\frac{k_{xp}^2}{j\omega\mu\epsilon_p}t_{mp}^{TM,A} \tag{2.55}$$

$$-\frac{jk_{xm}}{\mu}\left(1-r_{mp}^{TM,A}\right) = -\frac{jk_{xp}}{\mu}t_{mp}^{TM,A}. \tag{2.56}$$

Solving this equations for $r_{mp}^{TM,A}$ and $t_{mp}^{TM,A}$ gives:

$$r_{mp}^{TM,A}(k_z,\omega) = \frac{\epsilon_m k_{xp} - \epsilon_p k_{xm}}{\epsilon_m k_{xp} + \epsilon_p k_{xm}} \tag{2.57}$$

$$t_{mp}^{TM,A}(k_z,\omega) = \frac{2\epsilon_p k_{xm}^2}{k_{xp}(\epsilon_m k_{xp} + \epsilon_p k_{xm})}. \tag{2.58}$$

These coefficients are the ratios of $A_{z,r}$ and $A_{z,t}$ to A_{zi}. They may be converted to express the ratio of the electric fields, but the calculation is not simple because $\vec{E}$ contains two components for TM fields ($\hat{x}$ and $\hat{z}$). The reflection coefficients relating the amplitudes of the electric field are:

$$r_{mp}^{TM}(k_z,\omega) = \frac{n_p^2 k_{xm} - n_m^2 k_{xp}}{n_p^2 k_{xm} + n_m^2 k_{xp}} \tag{2.59}$$

$$t_{mp}^{TM}(k_z,\omega) = \frac{2n_m n_p k_{xm}}{n_p^2 k_{xm} + n_m^2 k_{xp}}. \tag{2.60}$$

Equations involving the reflection and transmission coefficients in other sources generally assume that the coefficients relate the amplitude of the electric field. For that reason, we will use the reflection and transmission coefficients of Equations 2.53-2.54 and 2.59-2.60. The conventions used for the direction of the electric and magnetic fields are shown in Figure 2.1.

If $n_1 > n_2$, then the incident wave can excite a field which is evanescent in the second medium. Because no power is transmitted into the second medium, the magnitude of the reflection coefficient must be 1. The reflection and transmission coefficients in this case are:

$$r_{mp}^{TE}(k_z,\omega) = \frac{jk_{zm} + \alpha_p}{jk_{zm} - \alpha_p} \tag{2.61}$$

$$t_{mp}^{TE}(k_z,\omega) = \frac{j2k_{zm}}{jk_{zm} - \alpha_p}, \tag{2.62}$$

where $\alpha_p = \sqrt{k_z^2 - k_p^2}$. Similarly, for TM polarization, the reflection and

transmission coefficients become:

$$r_{mp}^{TM}(k_z,\omega) = \frac{jn_p^2 k_{xm} + n_m^2 \alpha_{xp}}{jn_p^2 k_{xm} - n_m^2 \alpha_{xp}} \tag{2.63}$$

$$t_{mp}^{TM}(k_z,\omega) = \frac{j2n_m n_p k_{xm}}{jn_p^2 k_{xm} - n_m^2 \alpha_{xp}}. \tag{2.64}$$

If $e^{+j\omega t}$ time dependence is used, the complex conjugate must be taken of these expressions for both TE and TM polarizations.

2.3 Cylindrical Coordinates

2.3.1 TE^z and TM^z Fields

Solutions for TE^z and TM^z fields can be found using vector potentials of $\vec{F} = \hat{z}F_z$ and $\vec{A} = \hat{z}A$, respectively. Following Equation C.31 in Appendix C:

$$\frac{\partial^2 F_z}{\partial r^2} + \frac{1}{r}\frac{\partial F_z}{\partial r} + \frac{1}{r^2}\frac{\partial^2 F_z}{\partial \phi^2} + \frac{\partial^2 F_z}{\partial z^2} + k^2 F_z = 0. \tag{2.65}$$

In a uniform space, the separation of variables technique may be used, i.e. $F_z(r,\phi,z) = R(r)\Phi(\phi)Z(z)$. Substituting this into the above equation,

$$\Phi Z\frac{\partial^2 R}{\partial r^2} + \Phi Z\frac{1}{r}\frac{\partial R}{\partial r} + RZ\frac{1}{r^2}\frac{\partial^2 \Phi}{\partial \phi^2} + \Phi R\frac{\partial^2 Z}{\partial z^2} + k^2 R\Phi Z = 0. \tag{2.66}$$

Dividing this entire expression by $R\Phi Z$, we obtain:

$$\frac{1}{R}\frac{\partial^2 R}{\partial r^2} + \frac{1}{R}\frac{1}{r}\frac{\partial R}{\partial r} + \frac{1}{\Phi}\frac{1}{r^2}\frac{\partial^2 \Phi}{\partial \phi^2} + \frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} + k^2 = 0. \tag{2.67}$$

The second to last term on the left of the above expression is the only term dependent on z, and is a function of z alone. Similar to the reasoning near Equation 2.26, this term must be a constant, which will be defined as $-k_z^2$. The resulting differential equation is then:

$$\frac{1}{Z}\frac{\partial^2 Z}{\partial z^2} \equiv -k_z^2. \tag{2.68}$$

The general solution to which is

$$Z(z) = b_1 e^{jk_z z} + b_2 e^{-jk_z z}, \tag{2.69}$$

where b_1 and b_2 are constants which will be determined later by boundary conditions. Placing Equation 2.68 back into Equation 2.67 and multiplying the entire expression by r^2 yields

$$\frac{r^2}{R}\frac{\partial^2 R}{\partial r^2} + \frac{r}{R}\frac{\partial R}{\partial r} + (k^2 - k_z^2)r^2 + \frac{1}{\Phi}\frac{\partial^2 \Phi}{\partial \phi^2} = 0. \tag{2.70}$$

The last term in Equation 2.67 is a function of ϕ and is the only term in the equation to depend on ϕ. Therefore, it must be a constant, which we will define as $-m^2$.

$$\frac{1}{\Phi}\frac{\partial^2 \Phi}{\partial \phi^2} = -m^2. \tag{2.71}$$

The solution of this equation is

$$\Phi(\phi) = c_1 e^{jm\phi} + c_2 e^{-jm\phi}, \tag{2.72}$$

where c_1 and c_2 are constants determined by boundary conditions. Placing Equation 2.71 into Equation 2.70, we find:

$$r^2\frac{\partial^2 R}{\partial r^2} + r\frac{\partial R}{\partial r} + ((k^2 - k_z^2)r^2 - m^2)R = 0. \tag{2.73}$$

This is Bessel's differential equation, whose two linearly independent solutions are defined to the be the Bessel functions:

$$R(r) = a_1 J_m(k_r r) + a_2 Y_m(k_r r) \tag{2.74}$$

with $k_r \equiv \sqrt{k^2 - k_z^2}$.

A complete solutions for the TE^z and TM^z fields are, respectively, of the form:

$$F_z = (a_1 J_m(k_r r) + a_2 Y_m(k_r r))(b_1 e^{jk_z z} + b_2 e^{-jk_z z})(c_1 e^{jm\phi} + c_2 e^{-jm\phi}) \tag{2.75}$$

$$A_z = (a_1 J_m(k_r r) + a_2 Y_m(k_r r))(b_1 e^{jk_z z} + b_2 e^{-jk_z z})(c_1 e^{jm\phi} + c_2 e^{-jm\phi}). \tag{2.76}$$

2.3.2 Interpretation of the Solutions

Equation 2.74 is not the only form of the $R(r)$ solution. Because Maxwell's equations are linear, linear combinations of J_m and Y_m are themselves solutions. In particular, we choose

$$H_m^{(1)} \equiv J_m(x) + jY_m(x) \tag{2.77}$$

$$H_m^{(2)} \equiv J_m(x) - jY_m(x). \tag{2.78}$$

The resulting solution takes the form:

$$R(r) = \alpha_1 H_m^{(1)}(k_r r) + \alpha_2 H_m^{(2)}(k_r r). \tag{2.79}$$

The functions $H_m^{(1)}$ and $H_m^{(2)}$ are known as the Hankel functions. They may be interpreted as outwardly and inwardly propagating[2] waves, respectively.

Which form of the solution is appropriate to use depends on the details of the problem. J_m is finite at the origin while Y_m becomes infinite. For any physical solution, the field must be finite at all points, and therefore if the region of solution contains the origin, J_m must be the solution. In outer shells which do not contain the origin, it is convenient to use the Hankel functions, H_m^1 and H_m^2. Using J_m and Y_m in shell regions still theoretically results in the same solution; however, the interpretation as propagating waves is lost. From a computational perspective, there are advantages to maintaining the propagating wave interpretation.

The dispersion relationship in cylindrical coordinates is:

$$k_r^2 + k_z^2 = k^2, \tag{2.80}$$

which results from the definition of k_r in Equation 2.73. If either k_r^2 or k_z^2 is greater than k_z^2, it will force the other variable to be complex. If $k_z^2 < 0$, the situation is similar to that of planar coordinates. The field exponentially decays in the $\hat{z}$ direction. For $k_r^2 < 0$, many authors switch to modified Bessel functions.[3] However, this introduces needless complexity and special cases for the formulas, so in this work, regular Bessel and Hankel functions will be

[2] In some texts, $H_m^{(1)}$ is defined to be the inward-propagating wave and $H_m^{(2)}$ the outward-propagating wave. It is different here because of the $e^{-j\omega t}$ time dependence.

[3] $K_m(x) \equiv \frac{\pi}{2} j^{m+1} H^{(1)}(ix)$

Table 2.4: Electric and magnetic fields of TE^z solutions derived from the electric vector potential. Here B_m stands in for any one of the Bessel functions $J_m, K_m, H_m^{(1)}$, or $H_m^{(2)}$.

E_r	$-\frac{1}{\epsilon r}\frac{\partial F_z}{\partial \phi}$	$-\frac{jm}{\epsilon r}B_m(k_r r)e^{jk_z z}e^{jm\phi}$
E_ϕ	$\frac{1}{\epsilon}\frac{\partial F_z}{\partial r}$	$\frac{k_r}{\epsilon}B'_m(k_r r)e^{jk_z z}e^{jm\phi}$
E_z	0	0
H_r	$-\frac{1}{j\omega\mu\epsilon}\frac{\partial^2 F_z}{\partial r \partial z}$	$-\frac{k_r k_z}{\omega\mu\epsilon}B'_m(k_r r)e^{jk_z z}e^{jm\phi}$
H_ϕ	$-\frac{1}{j\omega\mu\epsilon r}\frac{\partial^2 F_z}{\partial \phi \partial z}$	$\frac{mk_z}{j\omega\mu\epsilon r}B_m(k_r r)e^{jk_z z}e^{jm\phi}$
H_z	$-\frac{1}{j\omega\mu\epsilon}\left(\frac{\partial^2}{\partial z^2}+k^2\right)F_z$	$-\frac{k_r^2}{j\omega\mu\epsilon}B_m(k_r r)e^{jk_z z}e^{jm\phi}$

Table 2.5: Electric and magnetic fields of TM^z solutions derived from the magnetic vector potential. Here B_m stands in for any one of the Bessel functions $J_m, K_m, H_m^{(1)}$, or $H_m^{(2)}$.

E_r	$-\frac{1}{j\omega\mu\epsilon}\frac{\partial^2 A_z}{\partial r \partial z}$	$-\frac{k_r k_z}{\omega\mu\epsilon}B'_m(k_r r)e^{jk_z z}e^{jm\phi}$
E_ϕ	$-\frac{1}{j\omega\mu\epsilon r}\frac{\partial^2 A_z}{\partial \phi \partial z}$	$\frac{mk_z}{j\omega\mu\epsilon r}B_m(k_r r)e^{jk_z z}e^{jm\phi}$
E_z	$-\frac{1}{j\omega\mu\epsilon}\left(\frac{\partial^2}{\partial z^2}+k^2\right)A_z$	$-\frac{k_r^2}{j\omega\mu\epsilon}B_m(k_r r)e^{jk_z z}e^{jm\phi}$
H_r	$\frac{1}{\mu r}\frac{\partial A_z}{\partial \phi}$	$\frac{jm}{\mu r}B_m(k_r r)e^{jk_z z}e^{jm\phi}$
H_ϕ	$-\frac{1}{\mu}\frac{\partial A_z}{\partial r}$	$-\frac{k_r}{\mu}B'_m(k_r r)e^{jk_z z}e^{jm\phi}$
H_z	0	0

used with a complex argument. With a complex argument, $H_m^{(1)}$ decays in the $+\hat{r}$ direction and $H_m^{(2)}$ decays in the $-\hat{r}$ direction, in analogy with the plane wave case.

Once again, $e^{jk_z z}$ represents a field propagating in the $+kz$ direction. The $e^{jm\phi}$ term represents a wave traveling around the origin counter-clockwise and the $e^{-m\phi}$ term represents a term traveling clockwise. The value of m must be an integer to ensure that the field is single-valued ($\Phi(\phi_0) = \Phi(\phi_0 + 2\pi)$).

2.3.3 Reflection Coefficients

The concept of reflection and transmission coefficients will now be extended to cylindrical coordinates. The physics is complicated by the fact that, unlike in Cartesian coordinates, in general cylindrical TE^z and TM^z modes mix at the interface between two materials. Because the modes mix, instead of two reflection coefficients (r_{TE} and r_{TM}), there will be four reflection reflec-

tion coefficients (r_{ee},r_{em},r_{me}, and r_{mm}).[4] The same is true for transmission coefficients.

Let there be a boundary between two materials at radius R. Let the wave be incident from medium 1 into medium 1, with refractive indices n_1 and n_2, respectively. By the phasematching argument presented in the plane wave section, both the reflected and transmitted waves must vary with the same k_z and m. The incident, reflected, and transmitted waves take the following form:

$$A_z = B_m(k_r r)e^{jk_z z}e^{jm\phi} \tag{2.81}$$

$$F_z = B_m(k_r r)e^{jk_z z}e^{jm\phi} \tag{2.82}$$

for TE and TM waves, respectively. The field components corresponding to these choices of F_z and A_z are shown in Tables 2.4 and 2.5, respectively.

To find the reflection and transmission coefficients, we must solve the following matrix equation expressing the continuity of E_z, H_z, E_ϕ, and H_ϕ:

$$\begin{bmatrix} 0 & 0 & \frac{k_{r1}^2}{j\omega\mu\epsilon_1}B_{m,r}(k_{r1}R) & -\frac{k_{r2}^2}{j\omega\mu\epsilon_2}B_{m,t}(k_{r2}R) \\ \frac{k_{r1}^2}{j\omega\mu\epsilon_1}B_{m,r}(k_{r1}R) & -\frac{k_{r2}^2}{j\omega\mu\epsilon_2}B_{m,t}(k_{r2}R) & 0 & 0 \\ -\frac{k_{r1}}{\epsilon_1}B'_{m,r}(k_{r1}R) & \frac{k_{r2}}{\epsilon_2}B'_{m,t}(k_{r2}R) & -\frac{mk_z}{j\omega\mu\epsilon_1 R}B_{m,r}(k_{r1}R) & \frac{mk_z}{j\omega\mu\epsilon_2 R}B_{m,t}(k_{r2}R) \\ -\frac{mk_z}{j\omega\mu\epsilon_1 R}B_{m,r}(k_{r1}R) & \frac{mk_z}{j\omega\mu\epsilon_2 R}B_{m,t}(k_{r2}R) & \frac{k_{r1}}{\mu}B'_{m,r}(k_{r1}R) & -\frac{k_{r2}}{\mu}B'_{m,t}(k_{r2}R) \end{bmatrix} \cdot\Gamma^{pol} = V^{pol} \tag{2.83}$$

where using

$$V^{TE} = \begin{bmatrix} 0 \\ -\frac{k_{r1}^2}{j\omega\mu\epsilon_1}B_{m,i}(k_{r1}R) \\ \frac{k_{r1}}{\epsilon_1}B'_{m,i}(k_{r1}R) \\ \frac{mk_z}{j\omega\mu\epsilon_1 R}B_{m,i}(k_{r1}R) \end{bmatrix} \tag{2.84}$$

generates $\Gamma^{TE} = [r_{ee}\ t_{ee}\ r_{em}\ t_{em}]^\top$, the reflection and transmission coefficients

[4]The naming convention used is that r_{em} represents is the reflection coefficient from a TE wave to a TM wave and so forth.

with an incident TE^z wave. Similarly, for TM^z waves, using

$$V^{TM} = \begin{bmatrix} -\frac{k_{r1}^2}{j\omega\mu\epsilon_1} B_{m,i}(k_{r1}R) \\ 0 \\ \frac{mk_z}{j\omega\mu\epsilon_1 R} B_{m,i}(k_{r1}R) \\ -\frac{k_{r1}}{\mu} B'_{m,i}(k_{r1}R)) \end{bmatrix} \tag{2.85}$$

generates the reflection and transmission coefficients $\Gamma^{TM} = [r_{me}\, t_{me}\, r_{mm}\, t_{mm}]^\top$. $B_{m,i}$, $B_{m,i}$, and $B_{m,i}$ are used as placeholders for the functions which are appropriate to a given situation. For example, in looking at the reflection of a wave propagating toward the origin, $B_{m,i}$ would be $H_m^{(2)}$, $B_{m,r}$ would be $H_m^{(1)}$, and $B_{m,t}$ would be $H_m^{(2)}$ (or J_m if medium 2 extended to contain the origin). Solutions for R_{EE}, R_{EM} obtained by solving this 4x4 matrix equation do not possess compact forms. It is generally more convenient to solve these equations numerically. The reflection and transmission coefficients are the ratios of the reflected or transmitted A_z and F_z fields to the incident A_z or F_z field.

A special case occurs if either $m = 0$ or $k_z = 0$. For TE^z modes, E_z and H_ϕ are zero in all layers. Similarly, E_ϕ and H_z fields are zero in all layers for TM^z modes. Looking at Equation 2.83, the problem may be separated into equations for TE and TM polarizations. That is, pure TE^z and TM^z modes are not mixed at boundaries if either $m = 0$ or $k_z = 0$. In this case, there are simple expressions for the reflection and transmission coefficients.

For TE^z fields, expressions for the continuity of the H_z and E_ϕ components of the magnetic and electric field are:

$$\frac{k_{r1}^2}{n_1^2}\left[B_{0,i}(k_{r1}R) + rB_{0,r}(k_{r1}R)\right] = t\frac{k_{r2}^2}{n_2^2}B_{0,t}(k_{r2}R) \tag{2.86}$$

$$\frac{k_{r1}}{n_1^2}\left[B'_{0,i}(k_{r1}R) + rB'_{0,r}(k_{r1}R)\right] = t\frac{k_{r2}}{n_2^2}B'_{0,t}(k_{r2}R). \tag{2.87}$$

When these equations are solved, the resulting reflection and transmission

coefficients are:

$$r^{TE} = -\frac{k_{r1}\frac{B_{0,i}(k_{r1}R)}{B'_{0,i}(k_{r1}R)} - k_{r2}\frac{B_{0,t}(k_{r2}R)}{B'_{0,t}(k_{r2}R)}}{k_{r1}\frac{B_{0,r}(k_{r1}R)}{B'_{0,i}(k_{r1}R)} - k_{r2}\frac{B_{0,t}(k_{r2}R)}{B'_{0,t}(k_{r2}R)}\frac{B'_{0,r}(k_{r1}R)}{B'_{0,i}(k_{r1}R)}} \tag{2.88}$$

$$t^{TE} = -\frac{n_2^2 k_{r1}}{n_1^2 k_{r2}}\frac{\frac{B_{0,i}(k_{r1}R)}{B'_{0,i}(k_{r1}R)} - \frac{B_{0,r}(k_{r1}R)}{B'_{0,r}(k_{r1}R)}}{k_{r1}\frac{B_{0,r}(k_{r1}R)}{B'_{0,r}(k_{r1}R)}\frac{B'_{0,t}(k_{r2}R)}{B'_{0,i}(k_{r1}R)} - k_{r2}\frac{B_{0,t}(k_{r2}R)}{B'_{0,i}(k_{r1}R)}}. \tag{2.89}$$

Similarly for TM^z fields, the equations expressing the continuity of the fields are:

$$\frac{k_{r1}^2}{n_1^2}\left[B_{0,i}(k_{r1}R) + rB_{0,r}(k_{r1}R)\right] = t\frac{k_{r2}^2}{n_2^2}B_{0,t}(k_{r2}R) \tag{2.90}$$

$$\frac{k_{r1}}{n_1^2}\left[B'_{0,i}(k_{r1}R) + rB'_{0,r}(k_{r1}R)\right] = t\frac{k_{r2}}{n_2^2}B'_{0,t}(k_{r2}R). \tag{2.91}$$

The corresponding reflection and transmission coefficients are:

$$r^{TM} = -\frac{n_2^2 k_{r1}\frac{B_{0,i}(k_{r1}R)}{B'_{0,i}(k_{r1}R)} - n_1^2 k_{r2}\frac{B_{0,t}(k_{r2}R)}{B'_{0,t}(k_{r2}R)}}{n_2^2 k_{r1}\frac{B_{0,r}(k_{r1}R)}{B'_{0,i}(k_{r1}R)} - n_1^2 k_{r2}\frac{B_{0,t}(k_{r2}R)}{B'_{0,t}(k_{r2}R)}\frac{B'_{0,r}(k_{r1}R)}{B'_{0,i}(k_{r1}R)}} \tag{2.92}$$

$$t^{TM} = -\frac{n_2^2 k_{r1}}{k_{r2}}\frac{\frac{B_{0,i}(k_{r1}R)}{B'_{0,i}(k_{r1}R)} - \frac{B_{0,r}(k_{r1}R)}{B'_{0,r}(k_{r1}R)}}{n_2^2 k_{r1}\frac{B_{0,r}(k_{r1}R)}{B'_{0,r}(k_{r1}R)}\frac{B'_{0,t}(k_{r2}R)}{B'_{0,i}(k_{r1}R)} - n_1^2 k_{r2}\frac{B_{0,t}(k_{r2}R)}{B'_{0,i}(k_{r1}R)}}. \tag{2.93}$$

2.4 Spherical Coordinates

Reflection and transmission coefficients will now be derived for spherical coordinates. The theory is very similar to that of cylindrical coordinates.

2.4.1 TE^r and TM^r Fields

Referring to Appendix C, the eigenequation for both TE and TM modes is:

$$\left(\nabla^2 + k^2\right)\psi = 0. \tag{2.94}$$

When the Laplacian is expanded in spherical coordinates, the resulting equation is:

$$\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial\psi}{\partial r}\right)+\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\psi}{\partial\theta}\right)+\frac{1}{r^2\sin^2\theta}\frac{\partial^2\psi}{\partial\phi^2}+k^2\psi=0. \quad (2.95)$$

Again, the technique of separation of variables will be invoked. Assume $\psi(r,\phi,\theta)=R(r)\Phi(\phi)\Theta(\theta)$. Placing this ansatz in Equation 2.95 and expanding yields:

$$\Phi\Theta\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+R\Phi\frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+R\Theta\frac{1}{r^2\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2}+k^2R\Phi\Theta=0. \quad (2.96)$$

Dividing through by $R(r)\Phi(\phi)\Theta(\theta)$ yields:

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+k^2r^2+\frac{1}{\Theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\frac{1}{\Phi}\frac{1}{\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2}=0. \quad (2.97)$$

The two terms on the left of the above equation depend only on r and are the only part of the equation dependent on r. By the familiar argument from Equation 2.26, they must, therefore, be equal to a constant. In anticipation of a future simplification, this constant will be defined as $\ell(\ell+1)$

$$\frac{1}{\Theta}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)+\frac{1}{\Phi}\frac{1}{\sin^2\theta}\frac{\partial^2\Phi}{\partial\phi^2}=\ell(\ell+1). \quad (2.98)$$

Multiplying through by $\sin^2\theta$ gives:

$$\frac{1}{\Theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)-\ell(\ell+1)\sin^2\theta+\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2}=0. \quad (2.99)$$

The rightmost term on the left side of the above equation depends only on ϕ and is the only term dependent on ϕ, so it again must be a constant. Choosing the value $-m^2$ for this constant, we obtain the following differential equation:

$$\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2}=-m^2, \quad (2.100)$$

which has the well-known solution

$$\Phi(\phi)=c_1e^{jm\phi}+c_2e^{-jm\phi}, \quad (2.101)$$

where c_1 and c_2 are constants to be determined by boundary conditions.

The remainder of Equation 2.99 is

$$\frac{1}{\Theta}\sin\theta\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right)-\ell(\ell+1)\sin^2\theta-m^2=0. \tag{2.102}$$

Now define $x=\cos\theta$ and $y(x)=\Theta(\theta(x))$. By making the proper substitutions[5] into the above equation, it may be transformed into

$$\frac{d}{dx}\left[\left(1-x^2\right)\frac{dy}{dx}\right]+\left[\ell(\ell+1)-\frac{m^2}{1-x^2}\right]y=0. \tag{2.103}$$

The above is the differential equation which defines the associated Legendre polynomials.

$$y(x)=b_1P_\ell^m(x)+b_2Q_\ell^m(x). \tag{2.104}$$

Then reversing the substitution, we recast the solution in terms of θ.

$$\Theta(\theta)=b_1P_\ell^m\left(\cos(\theta)\right)+b_2Q_\ell^m\left(\cos(\theta)\right), \tag{2.105}$$

where b_1 and b_2 are constants to be determined by boundary conditions.

Let us return to the left side of Equation 2.97. Recall the right side must be a constant $\ell(\ell+1)$. The equation obtained for the radial dependence is thus:

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right)+k^2r^2+\ell(\ell+1)=0. \tag{2.106}$$

Multiplying through by R and expanding the derivative on the left, we get:

$$r^2\frac{d^2R}{dr^2}+2r\frac{dR}{dr}+\left[k^2r^2+\ell(\ell+1)\right]R=0. \tag{2.107}$$

The solutions to this equation are defined to be the spherical Bessel functions:

$$R(r)=a_1j_\ell(kr)+a_2y_\ell(kr), \tag{2.108}$$

where a_1 and a_2 are constants set by boundary conditions.

[5] Use $\sin\theta=\sqrt{1-x^2}$ and $\frac{dx}{d\theta}=-\sin\theta$ to see that for an arbitrary function $f(x)$, $\frac{df}{d\theta}=\frac{df}{dx}\frac{dx}{d\theta}=-\sin\theta\frac{df}{dx}=-\sqrt{1-x^2}\frac{df}{dx}$.

The total solution for ψ is thus found to be:

$$\psi = [a_1 j_\ell(k_r r) + a_2 y_\ell(k_r r)] \cdot [b_1 P_\ell^m(\cos\theta) + b_2 Q_\ell^m(\cos\theta)] \cdot [c_1 e^{jm\phi} + c_2 e^{-jm\phi}]. \tag{2.109}$$

However, from Appendix C, the auxiliary electric and magnetic vector potentials are related to ψ by $F^r = r\psi_F$ and $A^r = r\psi_A$, respectively. So the total solution for the electric and magnetic vector potentials are:

$$F_r = r\,[a_1 J_\ell(k_r r) + a_2 Y_\ell(k_r r)] \cdot [b_1 P_\ell^m(\cos\theta) + b_2 Q_\ell^m(\cos\theta)] \cdot [c_1 e^{jm\phi} + c_2 e^{-jm\phi}] \tag{2.110}$$

$$A_r = r\,[a_1 J_\ell(k_r r) + a_2 Y_\ell(k_r r)] \cdot [b_1 P_\ell^m(\cos\theta) + b_2 Q_\ell^m(\cos\theta)] \cdot [c_1 e^{jm\phi} + c_2 e^{-jm\phi}]. \tag{2.111}$$

2.4.2 Interpretation of the Solutions

We begin interpreting Equations 2.110-2.111 by examining the angular components. Rodrigues's formula gives the ordinary Legendre polynomials:

$$P_\ell(x) = \frac{1}{2^\ell \ell!} \frac{d^\ell}{dx^\ell} \left(x^2 - 1\right)^\ell. \tag{2.112}$$

The associated Legendre polynomials of the first kind are:

$$P_\ell^m(x) = \frac{(-1)^m}{2^\ell \ell!} (1 - x^2)^{m/2} \frac{d^{\ell+m}}{dx^{\ell+m}} \left(x^2 - 1\right)^\ell. \tag{2.113}$$

It is clear from the derivative in the above equation that $P_\ell^m(x) = 0$ if $m > \ell$. Unlike Legendre polynomials of the first kind, Legendre polynomials of the second kind do not have a convenient analytical representation and can only be expressed as a power series. However, they do have the property that if $m \neq 0$, $Q_\ell^m(x) \to \infty$ as $x \to \pm 1$.

Because $Q_\ell^m(\cos\theta)$ become infinite at the poles (where $\cos\theta = 1$), if the region of solution contains either the positive or negative z axis, the solution must consist purely of Legendre polynomials of the second kind, $P_\ell^m(\cos\theta)$. These are the only types of solutions that will be considered in this work. Together, the angular components of the solution are the well-known spherical harmonic functions.

$$Y_\ell^m(\phi, \theta) = P_\ell^m(\cos\theta) e^{jm\phi}. \tag{2.114}$$

It can be shown that ℓ is the total angular momentum of the mode, while m is the projection of the angular momentum on the z axis. All the spherical harmonics with the same ℓ are in fact the same function, rotated about the origin.

For the radial portion of the solutions, it is convenient to work with the Riccatti-Bessel functions, defined as:

$$\hat{J}_\ell(x) = xj_\ell(x) = \sqrt{\frac{\pi x}{2}} J_{\ell+\frac{1}{2}}(x) \tag{2.115}$$

$$\hat{Y}_\ell(x) = -xy_\ell(x) = -\sqrt{\frac{\pi x}{2}} Y_{\ell+\frac{1}{2}}(x) \tag{2.116}$$

$$\hat{H}^{(1)}_\ell(x) = xh^{(1)}_\ell(x) = \sqrt{\frac{\pi x}{2}} H^{(1)}_{\ell+\frac{1}{2}}(x) \tag{2.117}$$

$$\hat{H}^{(2)}_\ell(x) = xh^{(2)}_\ell(x) = \sqrt{\frac{\pi x}{2}} H^{(2)}_{\ell+\frac{1}{2}}(x). \tag{2.118}$$

Depending on the author, $\hat{J}_m$ or $\hat{Y}_m$ may be referred to in the literature as S_m and C_m, respectively. With this definition, Equations 2.110 and 2.111 become:

$$F_r = \left[a_1\hat{J}_\ell(k_r r) + a_2\hat{Y}_\ell(k_r r)\right] \cdot [b_1 P^m_\ell \cos\theta) + b_2 Q^m_\ell(\cos\theta)] \cdot [c_1 e^{jm\phi} + c_2 e^{-jm\phi}] \tag{2.119}$$

$$A_r = \left[a_1\hat{J}_\ell(k_r r) + a_2\hat{Y}_\ell(k_r r)\right] \cdot [b_1 P^m_\ell \cos\theta) + b_2 Q^m_\ell(\cos\theta)] \cdot [c_1 e^{jm\phi} + c_2 e^{-jm\phi}]. \tag{2.120}$$

The interpretation of these functions is very similar to the cylindrical Bessel functions. $\hat{Y}_\ell$ becomes infinite at the origin while $\hat{J}_\ell$ does not. The functions $\hat{H}^{(1)}$ and $\hat{H}^{(2)}$ represent outwardly and inwardly traveling waves.

2.4.3 Reflection Coefficients in Spherical Coordinates

Reflection coefficients in spherical coordinates are very much like those in cylindrical coordinates with m=0. From phase matching considerations, ℓ and m must be conserved upon reflection or transmission. All of the fields,

Table 2.6: Components of TE^r field. Here $\hat{B}_\ell$ stands in for any one of the Riccati-Bessel functions $\hat{J}_m, \hat{K}_m, \hat{H}_m^{(1)}$, or $\hat{H}_m^{(2)}$.

E_r	0	0
E_θ	$-\frac{1}{\epsilon}\frac{1}{r\sin\theta}\frac{\partial F_r}{\partial\phi}$	$-\frac{jm}{\epsilon r\sin\theta}\hat{B}_\ell(kr)P_\ell^m(\cos\theta)e^{jm\phi}$
E_ϕ	$\frac{1}{\epsilon}\frac{1}{r}\frac{\partial F_r}{\partial\theta}$	$\frac{1}{\epsilon r}\hat{B}_\ell(kr)\frac{d}{d\theta}P_\ell^m(\cos\theta)e^{jm\phi}$
H_r	$-\frac{1}{j\omega\mu\epsilon}\left(\frac{\partial^2}{\partial r^2}+k^2\right)F_r$	$-\frac{k^2}{j\omega\mu\epsilon}\left(\hat{B}_\ell''(kr)+\hat{B}_\ell(kr)\right)P_\ell^m(\cos\theta)e^{jm\phi}$
H_θ	$-\frac{1}{j\omega\mu\epsilon}\frac{1}{r}\frac{\partial^2 F_r}{\partial r\partial\theta}$	$-\frac{k}{j\omega\mu\epsilon r}\hat{B}_\ell'(kr)\frac{d}{d\theta}P_\ell^m(\cos\theta)e^{jm\phi}$
H_ϕ	$-\frac{1}{j\omega\mu\epsilon}\frac{1}{r\sin\theta}\frac{\partial^2 F_r}{\partial r\partial\phi}$	$-\frac{mk}{\omega\mu\epsilon}\frac{1}{r\sin\theta}\hat{B}_\ell'(kr)P_\ell^m(\cos\theta)e^{jm\phi}$

Table 2.7: Electric and magnetic fields of TM^r solutions derived from the magnetic vector potential. Here $\hat{B}_\ell$ stands in for any one of the Riccati-Bessel functions $\hat{J}_m, \hat{K}_m, \hat{H}_m^{(1)}$, or $\hat{H}_m^{(2)}$.

E_r	$-\frac{1}{j\omega\mu\epsilon}\left(\frac{\partial^2}{\partial r^2}+k^2\right)A_r$	$-\frac{k^2}{j\omega\mu\epsilon}\left(\hat{B}_\ell''(kr)+\hat{B}_\ell(kr)\right)P_\ell^m(\cos\theta)e^{jm\phi}$
E_θ	$-\frac{1}{j\omega\mu\epsilon}\frac{1}{r}\frac{\partial^2 A_r}{\partial r\partial\theta}$	$-\frac{k}{j\omega\mu\epsilon r}\hat{B}_\ell'(kr)\frac{d}{d\theta}P_\ell^m(\cos\theta)e^{jm\phi}$
E_ϕ	$-\frac{1}{j\omega\mu\epsilon}\frac{1}{r\sin\theta}\frac{\partial^2 A_r}{\partial r\partial\phi}$	$-\frac{mk}{\omega\mu\epsilon}\frac{1}{r\sin\theta}\hat{B}_\ell'(kr)P_\ell^m(\cos\theta)e^{jm\phi}$
H_r	0	0
H_θ	$\frac{1}{\mu}\frac{1}{r\sin\theta}\frac{\partial A_r}{\partial\phi}$	$\frac{jm}{\mu r\sin\theta}\hat{B}_\ell(kr)P_\ell^m(\cos\theta)e^{jm\phi}$
H_ϕ	$-\frac{1}{\mu}\frac{1}{r}\frac{\partial A_r}{\partial\theta}$	$-\frac{1}{\mu r}\hat{B}_\ell(kr)\frac{d}{d\theta}P_\ell^m(\cos\theta)e^{jm\phi}$

incident, reflected, and transmitted, take the following form:

$$F_r = \hat{B}_\ell(kr)P_\ell^m(\cos\theta)e^{jm\phi} \tag{2.121}$$

$$A_r = \hat{B}_\ell(kr)P_\ell^m(\cos\theta)e^{jm\phi} \tag{2.122}$$

The field components corresponding to these two waves are given in Tables 2.6 and 2.7.

Consider the modes where $m = 0$. For these modes, the E_θ and H_ϕ components are zero in all layers for TE^r modes and the E_ϕ and H_θ components are zero in all layers for TM^r modes. As a result, TE and TM modes are not mixed by reflection or transmission at boundaries, and pure TE^r and TM^r modes exist. The TE reflection and transmission coefficients may be found by writing equations for the continuity of the tangential electric and

magnetic fields. The continuity equations for E_ϕ and H_θ are

$$\frac{1}{n_1^2}\left[\hat{B}_{\ell,i}(k_1R)+r^{TE}\hat{B}_{\ell,r}(k_1R)\right]=t^{TE}\frac{1}{n_2^2}\hat{B}_{\ell,t}(k_2R) \tag{2.123}$$

$$\frac{1}{n_1}\left[\hat{B}'_{\ell,i}(k_1R)+r^{TE}\hat{B}'_{\ell,r}(k_1R)\right]=t^{TE}\frac{1}{n_2}\hat{B}'_{\ell,t}(k_2R). \tag{2.124}$$

Solving these equations provides the reflection and transmission coefficients for TE^r modes, which are:

$$r^{TE}=-\frac{n_1\frac{\hat{B}_{\ell,t}(k_2R)}{\hat{B}'_{\ell,t}(k_2R)}-n_2\frac{\hat{B}_{\ell,i}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}}{n_1\frac{\hat{B}_{\ell,t}(k_2R)}{\hat{B}'_{\ell,t}(k_2R)}\frac{\hat{B}'_{\ell,r}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}-n_2\frac{\hat{B}_{\ell,r}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}} \tag{2.125}$$

$$t^{TE}=\frac{n_2}{n_1}\frac{\frac{\hat{B}_{\ell,i}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}-\frac{\hat{B}_{\ell,r}(k_1R)}{\hat{B}'_{\ell,r}(k_1R)}}{\frac{\hat{B}_{\ell,t}(k_2R)}{\hat{B}'_{\ell,i}(k_1R)}-\frac{\hat{B}_{\ell,r}(k_1R)}{\hat{B}'_{\ell,r}(k_1R)}\frac{\hat{B}'_{\ell,t}(k_2R)}{\hat{B}'_{\ell,i}(k_1R)}}. \tag{2.126}$$

Similarly for TM fields, the continuity equations for H_ϕ and E_θ are

$$\hat{B}_{\ell,i}(k_1R)+r^{TM}\hat{B}_{\ell,r}(k_1R)=t^{TM}\hat{B}_{\ell,t}(k_2R) \tag{2.127}$$

$$\frac{1}{n_1}\left[\hat{B}'_{\ell,i}(k_1R)+r^{TM}\hat{B}'_{\ell,r}(k_1R)\right]=t^{TM}\frac{1}{n_2}\hat{B}'_{\ell,t}(k_2R). \tag{2.128}$$

Solving for the reflection and transmission coefficients, the results are:

$$r^{TM}=-\frac{n_1\frac{\hat{B}_{\ell,i}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}-n_2\frac{\hat{B}_{\ell,t}(k_2R)}{\hat{B}'_{\ell,t}(k_2R)}}{n_1\frac{\hat{B}_{\ell,r}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}-n_2\frac{\hat{B}_{\ell,t}(k_2R)}{\hat{B}'_{\ell,t}(k_2R)}\frac{\hat{B}'_{\ell,r}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}} \tag{2.129}$$

$$t^{TM}=-\frac{n_2\frac{\hat{B}_{\ell,i}(k_1R)}{\hat{B}'_{\ell,i}(k_1R)}-n_2\frac{\hat{B}_{\ell,r}(k_1R)}{\hat{B}'_{\ell,r}(k_1R)}}{n_1\frac{\hat{B}_{\ell,r}(k_1R)}{\hat{B}'_{\ell,r}(k_1R)}\frac{\hat{B}'_{\ell,t}(k_2R)}{\hat{B}'_{\ell,i}(k_1R)}-n_2\frac{\hat{B}_{\ell,t}(k_2R)}{\hat{B}'_{\ell,i}(k_1R)}}. \tag{2.130}$$

$\hat{B}_{\ell,i}$, $\hat{B}_{\ell,r}$, and $\hat{B}_{\ell,t}$ are place holders for the actual functions which are appropriate to a given situation. For example, in looking at the reflection from an outer to an inner layer, $\hat{B}_{\ell,i}$ would be $\hat{H}_\ell^{(2)}$, $\hat{B}_{\ell,r}$ would be $\hat{H}_\ell^{(1)}$, and $\hat{B}_{\ell,t}$ would be $\hat{H}_\ell^{(2)}$ (or $\hat{J}_\ell$ if medium p was the innermost layer).

The reflection and transmission coefficients derived for the m=0 are in fact valid for any m. A simple argument of symmetry will prove this fact. The reflection and transmission coefficients should not change if the coordinate

system is rotated, as the physical wave is left unchanged. Recall that all $Y_\ell^m(\phi, \theta)$ functions are equivalent, but rotated. The coordinate system can then be rotated to transform in between them, and, therefore, the reflection coefficients must remain unchanged.

2.5 Conclusion

In this chapter, the reflection and transmission coefficients at single interfaces have been derived. In the next two chapters, it will be shown how to take those reflection coefficients for structures which contain multiple interfaces and use that knowledge to understand scattering, waveguides, and resonators.

CHAPTER 3

ANALYSIS OF SCATTERING

3.1 Introduction

Multilayer dielectric coatings are an essential component of modern optical technology. Dielectric mirrors with controlled dispersion characteristics must be carefully designed for femtosecond lasers [15, 16, 17]. Coating design techniques all go through a numerical optimization phase at some point in the algorithm, requiring repeated calculation of properties such as reflection, transmission, and dispersion [18, 19]. It is important in this design process that these calculations be fast and accurate.

The most common method for analyzing multilayer structures is the transfer matrix formalism [20] (also referred to as "transmission" or "ABCD" matrices in the literature). While in many cases it may accurately calculate the magnitude and phase of the reflection, finding phase derivatives requires the use of discrete derivative approximations, which can be both time-consuming and inaccurate. The transfer matrix formalism has been recast into a coupled-mode theory [21], which was later refined into a method which is both exact and scales linearly with the number of layers [22]. However, with any transfer matrix method, if loss is present or the field is evanescent in any layers, the transfer matrix problem may become ill-conditioned, resulting in a loss of accuracy.

In almost any introductory textbook on lasers, one may find an analysis of the transmission and reflection of light by a Fabry–Perot etalon based on Airy's analysis [11]. In these analyses the reflected field is written as a sum of multiply reflected amplitudes, each of which represents a portion of the field which has made a different number of round trips through the center medium. Previous work extending the reflection method to multilayer systems [23, 7] has resulted in expressions for the total reflectivity which

cannot be efficiently evaluated by computer and, in the case of Pai *et al.*, are not exact.

In this chpater it will be shown how to extend the Fabry–Perot analysis to multilayer structures and how to modify the technique to compute the phase derivatives of a structure exactly. The algorithm itself is analyzed, showing its computation complexity and the posedness of the problem. A software implementation of the new method is discussed.

3.2 Analysis

3.2.1 Mirror Reflectivity and Phase

Consider the three-layer dielectric structure shown in Figure 3.1a. The total reflected field is found by summing the geometric series:

$$\begin{aligned} r &= r_{12} + r_{23}t_{12}t_{12}e^{j2\phi}\left(1 + r_{21}r_{23}e^{j2\phi} + r_{21}^2 r_{23}^2 e^{j4\phi} + \ldots\right) \\ &= r_{12} + \frac{r_{23}t_{12}t_{21}e^{j2\phi}}{1 - r_{21}r_{23}e^{j2\phi}}, \end{aligned} \tag{3.1}$$

where $r_{mp} = \frac{n_m - n_p}{n_m + n_p}$, $t_{mp} = \frac{2n_m}{n_p + n_m}$, and ϕ is the phase accumulated propagating through the center layer.

The above analysis is difficult to directly extend to structures containing more than three layers because it is necessary to sum over every possible path. To keep track of the reflections, a state diagram may be constructed, shown in Figure 3.1b. The diagram shows the amplitude and phase accumulated as the field propagates to neighboring points. These amplitudes and phases result from reflection, transmission, and propagation. The state diagram may then be translated into a transition matrix, which relates the field at a set of points in the mirror with the field at those same points after the light has undergone a reflection or transmission. For the one-layer structure with

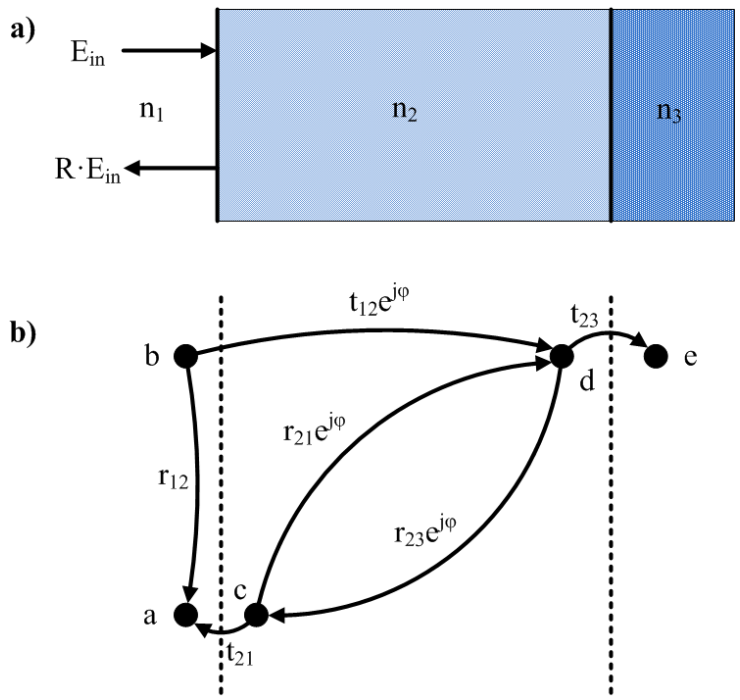

Figure 3.1: (a) Physical picture of three layer dielectric structure (b) Equivalent state diagram.

states labeled as shown in Figure 3.1b, the transition matrix is:

$$T = \begin{pmatrix} 1 & r_{12} & t_{21} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_{23}e^{j\phi_1} & 0 \\ 0 & t_{12}e^{j\phi_1} & r_{21}e^{j\phi_1} & 0 & 0 \\ 0 & 0 & 0 & t_{23} & 1 \end{pmatrix}. \tag{3.2}$$

Note that the exact positions of the points of the state diagram within the layers are irrelevant. For convenience, the positions shown were chosen so that T may be written as a product of diagonal matrix, containing the phase terms, and a sparse matrix, which comprises the transmission and reflection coefficients. Most of the frequency dependence comes from the variation of the phase terms, while the reflection and transmission coefficients are weakly dependent on frequency. Additional states may also be inserted anywhere in the structure without affecting the calculated field at the original points.

The transition matrix analysis is highly analogous to the theory of Markov chains (see Appendix A), from which the notation and several theorems used

in this analysis are taken. In Markov language, states 1 and 5 of Figure 3.1 are known as absorbing states because after the field enters one of them, it may not return to any other states. The remaining states are known as transient states. The field in the transient states will eventually go to zero after a large number of reflections, as all the field accumulates in the absorbing states.

Consider the submatrix formed by transitions among the transient states. This matrix will be labeled the Q matrix. For the diagram of Figure 3.1, the resulting matrix is:

$$Q = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & r_{23}e^{j\phi_1} \\ t_{12}e^{j\phi_1} & r_{21}e^{j\phi_1} & 0 \end{pmatrix}. \tag{3.3}$$

For the purpose of calculating mirror reflectivity, the mirror will be excited by the rightward traveling wave immediately outside the first dielectric layer. For simplicity, a unit excitation may be assumed $E_0 = (1, 0, 0...0)^T$. The column vector begins with state 2 because only the transient states are considered. The field after one step through the system is:

$$E_1 = QE_0 \tag{3.4}$$

and in general, the field vector after k transitions is:

$$E_k = Q^k E_0. \tag{3.5}$$

The total field vector is the sum of the field after every step:

$$\begin{aligned} E^{(0)} &= E_0 + E_1 + E_2 + \ldots \\ &= \left(IE_0 + QE_0 + Q^2E_0 + \ldots\right) \\ &= \left(I + Q + Q^2 + \ldots\right) E_0 \\ &= (I - Q)^{-1} E_0 \\ &= NE_0, \end{aligned} \tag{3.6}$$

where $N \equiv (I - Q)^{-1}$. The N matrix has the interpretation that the N_{mp} element is the field at point m if the structure is excited with unit amplitude at point p. The complex column vector $E^{(0)}$ is interpreted to be the strength of the field at all of the points in the state diagram in Figure 3.1.

The terms in the T matrix which couple the transient states to the absorbing states are referred to as the coupling matrix. In the example used so far, the coupling matrix is:

$$C = \begin{pmatrix} r_{12} & t_{21} & 0 \\ 0 & 0 & t_{23} \end{pmatrix} \tag{3.7}$$

The field at the two absorbing states of the mirror is:

$$E_{rt} = CNE_0, \tag{3.8}$$

where E_{rt} is a column vector comprising the reflected and transmitted fields.

This formalism may be simplified by adding two imaginary layers, one on each end of the structure and with the same refractive index as the surrounding media, as shown in Figure 3.2. Physically, the layers have no effect because they have zero thickness. Mathematically, these two layers have the effect of reducing the coupling matrix to

$$C = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}. \tag{3.9}$$

as $r_{11} = 0$ and $t_{11} = t_{55} = 1$. Pairs of states in Figure 3.2b are highlighted by having red, green, orange, and blue outlines. Each of these pairs of states is separated by zero distance and an interface with no change in the refractive index. The electric fields are therefore identical for each pair of points. Effectively, adding the zero-thickness layers created two transient states with the same electric field as the absorbing states. The reflection and transmission coefficients can therefore be read directly from the corresponding elements of the $E^{(0)}$ vector.

$$\bar{R}^{(0)} = E_2^{(0)} \tag{3.10}$$

$$\bar{T}^{(0)} = E_{end}^{(0)}. \tag{3.11}$$

The subscripts index the vector $E^{(0)}$ (starting from 1).

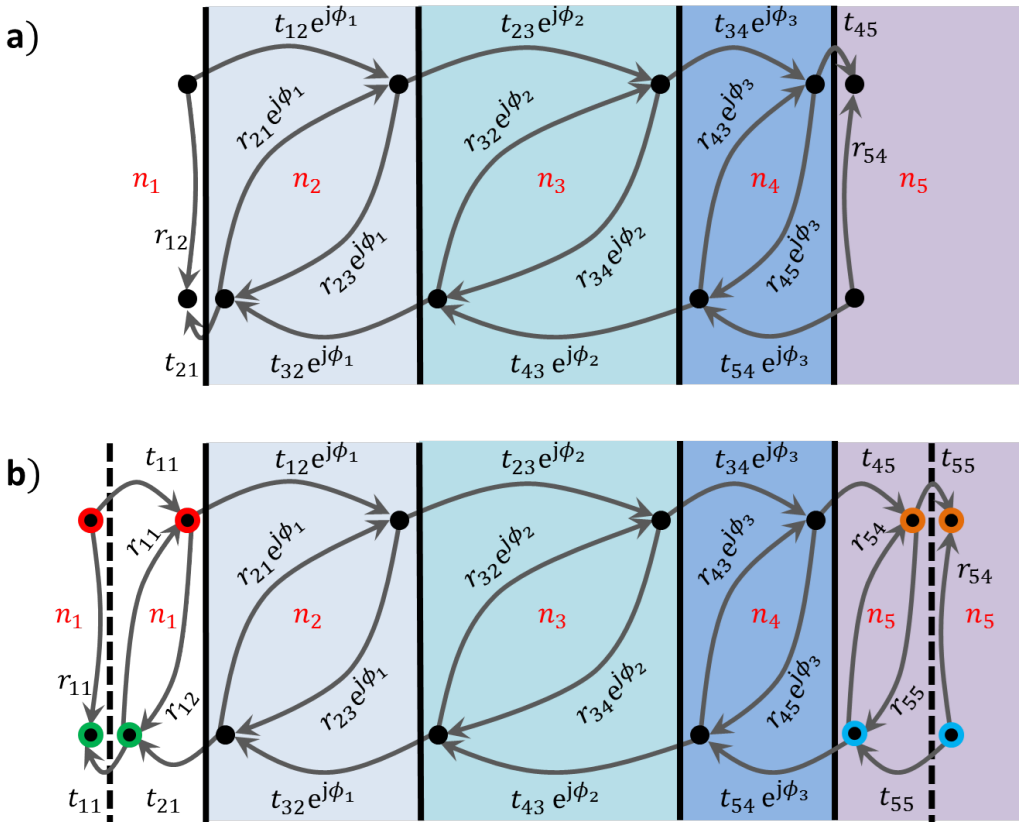

Figure 3.2: (a) Three-layer mirror structure with state diagram superimposed on top of it. (b) Same mirror structure with zero-thickness layers (indicated by the dashed lines) inserted outside the front and rear boundary of the mirror. Because the change in refractive index is zero, $t_{11} = t_{55} = 1$ and $r_{11} = r_{55} = 0$.

3.2.2 Mirror Dispersion

In applications involving short pulses, the second and higher-order derivatives of the phase, which primarily determine how a mirror distorts pulses [24], are as important as the magnitude of the reflectivity itself. The primary source of dispersion in a dielectric mirror is the structure of the mirror (layer thicknesses and indices of refraction). In addition, the materials which comprise the mirror add dispersion. The following analysis will include both effects.

For this discussion, it will be assumed that index-matched, zero-thickness layers, discussed at the end of the last section, have been inserted. This allows one to read the reflection and transmission coefficients directly from

$E^{(0)}$. Recall the vector of electric fields is given by:

$$E^{(0)} = NE_0. \tag{3.12}$$

Taking derivatives of the above equation is the first step to finding the mirror dispersion. The vector of the incident field, E_0, is a constant with respect to frequency and therefore its derivative is zero. Then we have only to find the derivatives of N. Using the formula $\frac{d}{dx}A^{-1}(x) = A^{-1}\frac{dA}{dx}A^{-1}$, the first, second, and third derivatives are:

$$\begin{aligned} E^{(1)} &= NQ^{(1)}NE_0 \\ &= NQ^{(1)}E^{(0)} \qquad (3.13) \\ E^{(2)} &= NQ^{(2)}E^{(0)} + 2NQ^{(1)}NQ^{(1)}E^{(0)} \\ &= NQ^{(2)}E^{(0)} + 2NQ^{(1)}E^{(1)} \qquad (3.14) \\ E^{(3)} &= NQ^{(3)}NE_0 + 3NQ^{(2)}NQ^{(1)}NE_0 \\ &\quad +3NQ^{(1)}\left(2NQ^{(1)}E^{(1)} + NQ^{(1)}NQ^{(1)}E^{(0)}\right) \\ &= NQ^{(3)}E^{(0)} + 3NQ^{(2)}E^{(1)} + 3NQ^{(1)}E^{(2)}, \qquad (3.15) \end{aligned}$$

where Equations 3.13, 3.14, and 3.15 have been rearranged to avoid redundant calculations.

The derivatives of the matrix Q are simply the derivatives of the individual elements of the matrix. Every nonzero element of the Q matrix takes the form $r_{mp}e^{j\phi_{mp}}$ or $t_{mp}e^{j\phi_{mp}}$, depending on whether the term is a reflection or transmission term, respectively. Without loss of generality, consider a reflective term, such as

$$Q_{mp} = r_{mp}e^{j\phi_{mp}}. \tag{3.16}$$

The derivatives of this term are:

$$\begin{aligned} Q^{(1)}_{mp} &= e^{j\phi_{mp}}\left[r'_{mp} + j\phi'_{mp}r_{mp}\right] \qquad (3.17) \\ Q^{(2)}_{mp} &= e^{j\phi_{mp}}\left[r''_{mp} + j2\phi'_{mp}r'_{mp} + (j\phi''_{mp} - \phi'^2_{mp})r_{mp}\right] \qquad (3.18) \\ Q^{(3)}_{mp} &= e^{j\phi_{mp}}\left[r'''_{mp} + j3\phi'_{mp}r''_{mp} + 3(j\phi''_{mp} - \phi'^2_{mp})r'_{mp}\right. \\ &\quad \left.+(j\phi'''_{mp} - j\phi'^3_{mp} - 3\phi'_{mp}\phi''_{mp})r_{mp}\right]. \qquad (3.19) \end{aligned}$$

where the terms have been grouped by derivatives of r_{mp}.

The derivatives of the reflection and transmission coefficients, valid at any angle of incidence, will be discussed in Section 3.2.4. Here the derivatives of the phase at normal incidence will be discussed. The phase of propagation of the mp^{th} element of the matrix is

$$\phi_{mp}(\omega) = \frac{\omega n_{mp}(\omega) L_{mp}}{c}, \tag{3.20}$$

where ω is the angular frequency of the light, L_{mp} is the thickness of the layer, c is the speed of light in vacuum, and the index of refraction has been written $n_{mp}(\omega)$ to emphasize that is a function of frequency. Taking derivatives of Equation 3.20, one finds:

$$\phi'_{mp} = \frac{L_{mp}}{c}\left(n_{mp} + \omega\frac{dn_{mp}}{d\omega}\right) \tag{3.21}$$

$$\phi''_{mp} = \frac{L_{mp}}{c}\left(2\frac{dn_{mp}}{d\omega} + \omega\frac{d^2n_{mp}}{d\omega^2}\right) \tag{3.22}$$

$$\phi'''_{mp} = \frac{L_{mp}}{c}\left(3\frac{d^2n_{mp}}{d\omega^2} + \omega\frac{d^3n_{mp}}{d\omega^3}\right) \tag{3.23}$$

and, in general,

$$\phi^{(q)}_{mp} = \frac{L_{mp}}{c}\left(q\frac{d^{q-1}n}{d\omega^{q-1}} + \omega\frac{d^qn}{d\omega^q}\right). \tag{3.24}$$

The functional dependence of the phase on ω comes primarily from the change in wavelength with respect to frequency, which corresponds to the left term in Equation 3.21. The right term corresponds to the change in index of refraction with frequency, known as dispersion, which is a much smaller effect. Dispersion also affects the derivatives of r_{mp} and t_{mp}, as will discussed in Section 3.2.4. This is an even smaller effect.

As shown in Appendix B, the phase derivatives can be found from the

total derivatives by:

$$r_\omega = \bar{R}^{(1)}_{\omega,\hat{r}} \tag{3.25}$$

$$\theta_\omega = \frac{1}{r}\bar{R}^{(1)}_{\omega,\hat{\theta}} \tag{3.26}$$

$$r_{\omega\omega} = \bar{R}^{(2)}_{\hat{r}} + r\theta^2_\omega \tag{3.27}$$

$$\theta_{\omega\omega} = \frac{1}{r}\left(\bar{R}^{(2)}_{\omega\omega,\hat{\theta}} - 2r_\omega\theta_\omega\right) \tag{3.28}$$

$$\theta_{\omega\omega\omega} = \frac{1}{r}\left(\bar{R}^{(3)}_{\omega\omega\omega,\hat{\theta}} - 3r_\omega\theta_{\omega\omega} - 3r_{\omega\omega}\theta_\omega + r\theta^3_\omega\right), \tag{3.29}$$

where $\bar{R}^{(m)}_\rho$ and $\bar{R}^{(m)}_\theta$ are the parallel and perpendicular projections of the change in reflectivity on the unperturbed reflectivity. Equations 3.26, 3.28, and 3.29 represent exact formulas for the first three phase derivatives of the mirror reflectivity.

3.2.3 Optimized Calculation

While the method described in the past two sections produces correct results, it inefficiently requires calculating the full inverse of $(I-Q)$. In this section, it will be shown how to translate the formulas of the last two sections into the form $Ax = b$, where A is a matrix and x and b are column vectors. Equations of this form may be solved with sparse matrix methods, without the need to explicitly calculate any inverse, significantly increasing the speed of the calculation.

Recall Equation 3.6, where it was found that $E^{(0)} = NE_0 = (I-Q)^{-1}E_0$. This equation can be altered by multiplying by $N^{-1} = (I-Q)$ so that

$$(I-Q)E^{(0)} = E_0 \tag{3.30}$$

is obtained. Here E_0 (the input field) is known and $E^{(0)}$ is unknown. Similarly, Equations 3.13-3.15 can be rearranged to yield:

$$(I-Q)E^{(1)} = \Delta Q^{(1)}E^{(0)} \tag{3.31}$$

$$(I-Q)E^{(2)} = \Delta Q^{(2)}E^{(0)} + 2\Delta Q^{(1)}E^{(1)} \tag{3.32}$$

$$(I-Q)E^{(3)} = \Delta Q^{(3)}E^{(0)} + 3\Delta Q^{(2)}E^{(1)} + 3\Delta Q^{(1)}E^{(2)}. \tag{3.33}$$

Because of the sparse, nearly diagonal nature of $I-Q$, these linear equations can be solved efficiently with a sparse matrix solver.

3.2.4 Reflectivity at Non-Normal Incidence

Consider a plane wave incident on a dielectric stack at an angle. Compared with the theory reflection at normal incidence, there are two changes which must be taken into account compared to the theory for reflection at normal incidence: First, the reflection and transmission at each layer are altered, and second, the phase from propagating through each layer changes.

Let the $\hat{x}-\hat{z}$ plane be the plane of incidence and without loss of generality, let $\hat{x}$ be perpendicular to the surface of the mirror. The angle of incidence, θ, is specified in the first medium. It will be more convenient to work with $k_{z0} = n_0 \sin(\theta)$, rather than θ. Due to phase-matching considerations, k_{z0} must be the same in all layers. Fresnel's equations for reflection and transmission coefficients coefficients for $\hat{s}$ polarized light are:

$$r_{mp}^{TE}(k_z,\omega) = \frac{k_{xm}-k_{xp}}{k_{xm}+k_{xp}} \tag{3.34}$$

$$t_{mp}^{TE}(k_z,\omega) = \frac{2k_{xm}}{k_{xm}+k_{xp}}, \tag{3.35}$$

while the corresponding coefficients for $\hat{p}$ polarized light are:

$$r_{mp}^{TM}(k_z,\omega) = \frac{n_m^2 k_{xp}-n_p^2 k_{xm}}{n_p^2 k_{xm}+n_m^2 k_{xp}} \tag{3.36}$$

$$t_{mp}^{TM}(k_z,\omega) = \frac{2n_m n_p k_{xm}}{n_p^2 k_{xm}+n_m^2 k_{xp}}, \tag{3.37}$$

where m indexes each layer and $k_{xm} = \sqrt{k_m^2 - k_{z0}}$. This accounts for the change in reflectivity at each interface with respect to angle of incidence. Next we must account for how the phase of each layer changes at various angles of incidence. The phase of propagating through the m^{th} layer is:

$$\begin{aligned}\phi_m &= L_m k_{xm} \\ &= L_m\sqrt{k_m^2 - k_{z0}^2}.\end{aligned} \tag{3.38}$$

The derivatives of the reflection and transmission coefficients and phase

will now be discussed. These quantities are necessary to compute Equations 3.17-3.19. For the reflection coefficients, taking analytical derivatives of Equations 3.34-3.37 results in very long expressions. For computational purposes, it is convenient to define f as the numerator and g as the denominator (whose derivatives are obvious) and the combine the expressions to get the total derivatives with the following expressions:

$$r_{mp} = \frac{f}{g} \tag{3.39}$$

$$r'_{mp} = \frac{f'}{g} - \frac{fg'}{g^2} \tag{3.40}$$

$$r''_{mp} = \frac{f''}{g} - \frac{2f'g' + fg''}{g^2} + \frac{2fg'^2}{g^3} \tag{3.41}$$

$$r'''_{mp} = \frac{f'''}{g} - \frac{fg''' + 3f'g'' + 3f''g'}{g^2} + \frac{6f'g'^2 + 6fg'g''}{g^3} - \frac{6fg'^3}{g^4}. \tag{3.42}$$

The derivatives of Equation 3.38 are now taken with respect to frequency.

$$\gamma'_m = \frac{d\phi_m}{d\omega} = L_m \frac{k_m k'_m - k_{z0} k'_{z0}}{(k_m^2 - k_{z0}^2)^{1/2}} \tag{3.43}$$

$$\begin{aligned}\gamma''_m = \frac{d^2\phi_m}{d\omega^2} = L_m \big[&k_m^3 k''_m - k_{z0}^2 k'^2_m - k_{z0} k''_{z0} k_m^2 + 2k_{z0} k'_{z0} k_m k'_m - k_{z0}^2 k_m k''_m \\ &- k'^2_{z0} k_m^2 + k_{z0}^3 k''_{z0}\big] / (k_m^2 - k_{z0}^2)^{3/2}\end{aligned} \tag{3.44}$$

$$\begin{aligned}\gamma'''_m = \frac{d^3\phi_m}{d\omega^3} = L_m \big[&3(k_m k'_m - k_{z0} k'_{z0})^3 \\ &- 3(k_m^2 - k_{z0}^2)(k_m k'_m - k_{z0} k'_{z0})(k_m k''_m + k'^2_m - k_{z0} k''_{z0} - k'^2_{z0}) \\ &+ (k_m^2 - k_{z0}^2)^2 (k_m k'''_m + 3k'_m k''_m - k_{z0} k'''_{z0} - 3k'_{z0} k''_{z0})\big] / (k_m^2 - k_{z0}^2)^{5/2},\end{aligned} \tag{3.45}$$

where k_m, k'_m, k'_m, and $k_m^{(3)}$ are defined in Equations 3.20-3.23. If the γ_m's from Equations 3.38-3.45 replace the ϕ_{mp}'s in Equations 3.17-3.19, then the phase derivatives at non-normal incidence may be obtained. There is a subtlety in using these equations: If dispersion is to be simulated at a fixed angle of incidence, then the derivatives of k_{z0} must be retained, as the wavevector and its projection onto the $\hat{z}$-axis increases with frequency. However, if dispersion is to be calculated at fixed k_{z0} (such as in calculations for a waveguide structure), then all the derivatives of k_{z0} should be set to zero.

3.2.5 Lossy Interfaces, Lossy Materials, and Anisotropic Media

Scattering at interfaces is easily introduced into the Markov formalism. Nowhere in the derivations above was it required that r_{ij} and t_{ij} be subject to the constraint: $1 + r_{mp} = t_{mp}$, i.e. that all of the energy of an incident plane wave is reflected or transmitted as plane waves. Likewise loss materials with a complex index of refraction are easily introduced into the Markov formalism and the appropriate reflection and transmission coefficients. Finally, anisotropic media may be included by adding two additional states per layer. The reflection and transmission coefficients must be changed. The polarization states will be coupled by propagation through the anisotropic layers.

3.3 Computational Results and Discussion

3.3.1 Computational Complexity

In this section the run time complexity of the algorithm will be examined. For this section, let n be equal to the total number of layers in the mirror structure and d be the order of the largest derivative to be calculated. The matrix $(I - Q)$ is a square, sparse matrix with a side dimension of $2n + 1$. All of its nonzero terms are contained within a pentadiagonal matrix. Hence the number of nonzero terms is linear in the number of mirror layers

First, consider the naïve Markov method, in which the full N matrix must be computed. Because the starting matrix (I-Q) is sparse, the inversion process requires $\mathcal{O}(n^2)$ floating point operations [25]. To find the d^{th} derivative, sums such as those in Equations 3.13-3.15 must be computed. The number of terms in the sum is linear in the order of the derivative to be computed, but all the $d - 1$ derivatives must be known, and because is matrix is full, each multiplication takes $\mathcal{O}(n^2)$ time. The total complexity is therefore $\mathcal{O}(n^2 d^2)$.

For the improved algorithm presented in Section 3.2.3, one simply needs to solve Equation 3.30, where E_0 is known and E is unknown, which requires $\mathcal{O}(n)$ time [25]. Computing the d^{th} derivative involves computing sums like those seen on the right sides of Equations 3.32 and 3.33. The number of terms in these sums is linear in the order of the derivative being calculated

and requires knowledge of the last $d-1$ derivatives. In this case the matrices are sparse, however, so that matrix multiplication and addition may be accomplished in linear time. The computation complexity is therefore $\mathcal{O}(nd^2)$. For computing derivatives fourth order and lower, it has been found that the time taken in solving the linear system overwhelmingly dominates the time necessary to multiply the matrices. The solution of one linear system is required per order of derivative. For low order derivatives then, the computational complexity scales as $\mathcal{O}(nd)$.

3.3.2 Implementation and Validation

The Markov algorithm was implemented in MATLAB. It utilizes the built-in sparse matrix solver, *mldivide*. The code is available at Matlab Central.[1] In order to examine the effects of various approximations, the program is capable of calculating phase derivatives of up to third order exactly, including the derivatives of the reflection and transmission coefficients up to third order. The implementation was validated by comparing output from the program with analytical results, specifically, the predictions of the exact expression in [20] for the reflection from a finitely repeating, two-layer structure. Comparisons between the two were extensive and have included variations in layer thicknesses, refractive indices, angle of incidence, and polarization of the optical excitation field.

Representative results are illustrated in Figure 3.3 for a periodic mirror structure comprising 100 pairs of SiO_2-TiO_2 layers in which the SiO_2 and TiO_2 layer thicknesses are 185 nm and 70 nm, respectively. This entire mirror is assumed to be bounded by SiO_2. Light is incident at 45° with p-polarization. Panels (a), (b), and (c) of Figure 3.3 present the relative error, in the 650-800 nm wavelength region, between the first, second, and third phase derivatives (respectively) calculated by the Markov-Airy formalism and those given by the exact expression given in [20] for the reflectivity of an alternating two-layer mirror structure. It is evident that, for the first phase derivative (Figure 3.3(a)), the difference between the result calculated by the Markov method and that given by the analytic expression [20] is near the limit of the precision of the machine. However, as the order of the phase

[1] http://www.mathworks.com/matlabcentral/fileexchange/47360-markov-airy-mirror-zip

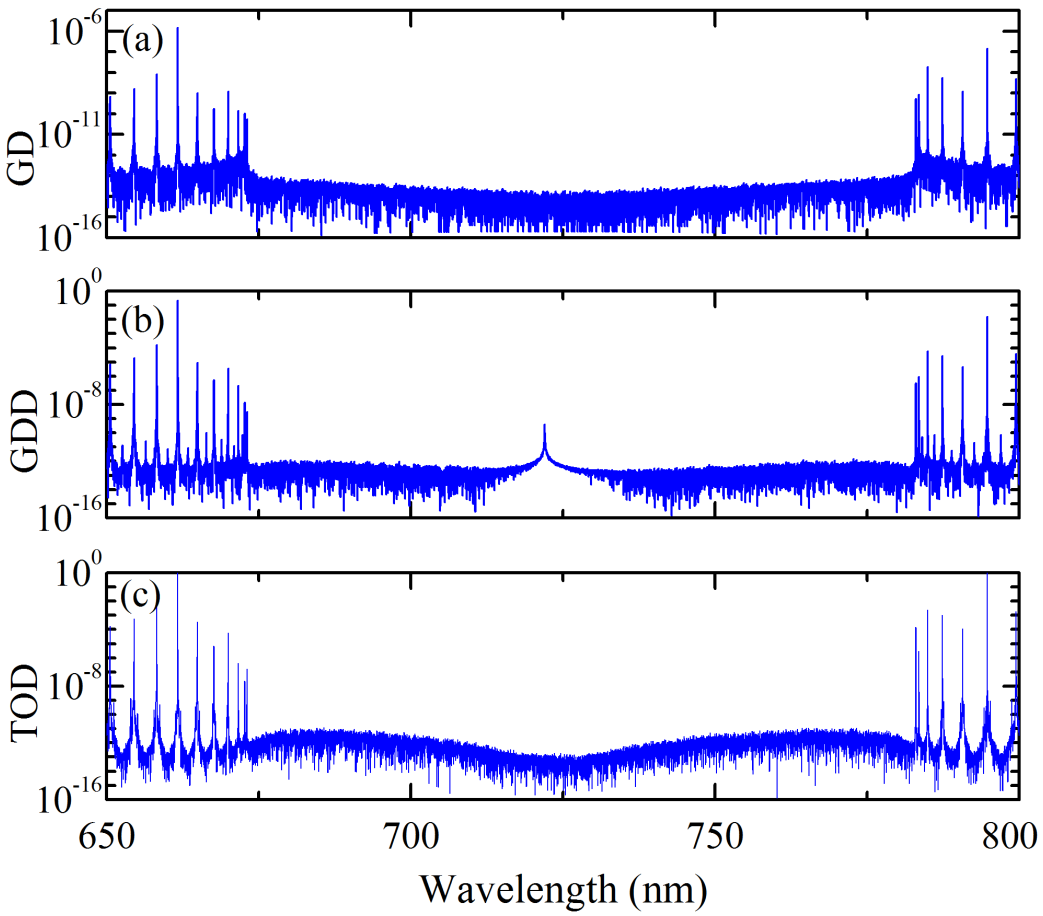

Figure 3.3: Relative error between the phase derivatives calculated by the Markov-Airy formalism and an analytic expression [20]. Results are given for the: (a) first, (b) second, and (c) third phase derivatives, and a mirror structure consisting of 100 pairs of SiO_2/TiO_2 flims. The thicknesses of the SiO_2 and TiO_2 layers are 185 nm and 70 nm, respectively, and the entire mirror structure is assumed to be bounded by SiO_2.

derivative increases, the error associated with the difference between the two theoretical formulations rises. The spikes which appear in the relative error in Figure 3.3 occur at wavelengths where the exact value passes through zero.

More comprehensive tests of the Markov-Airy formalism were undertaken with the goal of evaluating several approximations commonly adopted in the literature when phase derivatives are calculated for an ultra-broadband mirror. As one example, the design for a SiO_2-Nb_2O_5, dispersion-compensated mirror of [26] was adopted and the material parameters were assumed to be those of [27]. Figure 3.4 plots the relative errors ($|\frac{(approx-exact)}{exact}|$) in the phase derivatives of under various approximations. Results are presented here for

the 700-900 nm wavelength interval at normal incidence. For each of the lowest three phase derivatives (Figure 3.4(a)-(c) for the first, second, and third derivatives, respectively), the effect on the relative error of various assumptions regarding the wavelength dependence of the refractive index (n) or the reflectivity (r) of an individual mirror layer is illustrated by a separate curve. In the figure, the first, second and third derivatives of n with respect to optical frequency are denoted as n', n'' , and n''', respectively, and when one derivative (such as n'') is set to zero, all higher orders are neglected as well. Similar comments apply to the derivatives of r.

One notices immediately in Figure 3.4 that ignoring the derivatives of n to the same order as the order of the phase derivative in question results in an unacceptably large error (10^{-2}). Furthermore, neglecting r' in the calculations is reasonable when determining the first and second phase derivatives (Figure 3.4(a) and (b), respectively), but retaining r' is essential when calculating the third order phase derivative. Similarly, the penalty resulting from ignoring the second and third derivatives of r is generally small, typically leading to a relative error of $<10^{-3}$.

A more severe impact of neglecting higher derivatives of r is observed if a broadband mirror has fewer layers. Consider, for example, a mirror comprising a SiO_2-TiO_2 quarter wave stack. If the angle of incidence is assumed to be 45 degrees with s-polarization, the mirror has 3 layers (SiO_2-TiO_2-SiO_2), and the central wavelength for the mirror's reflectivity is chosen to be 800 nm, then the influence of the r'' and r''' on the accuracy of the calculations is magnified. Because this mirror is shorter than that of Figure 3.4, a greater fraction of the dispersion supplied by the mirror is attributable to the frequency-dependent characteristics of optical reflection and transmission at the layer interfaces. This conclusion should be borne in mind when examining thin optical structures and, in particular, calculating the group velocity and group velocity dispersion for a planar waveguide. It should also be noted from Figures 3.4 and 3.5 that neglecting any of the derivatives of n leads to large relative errors.

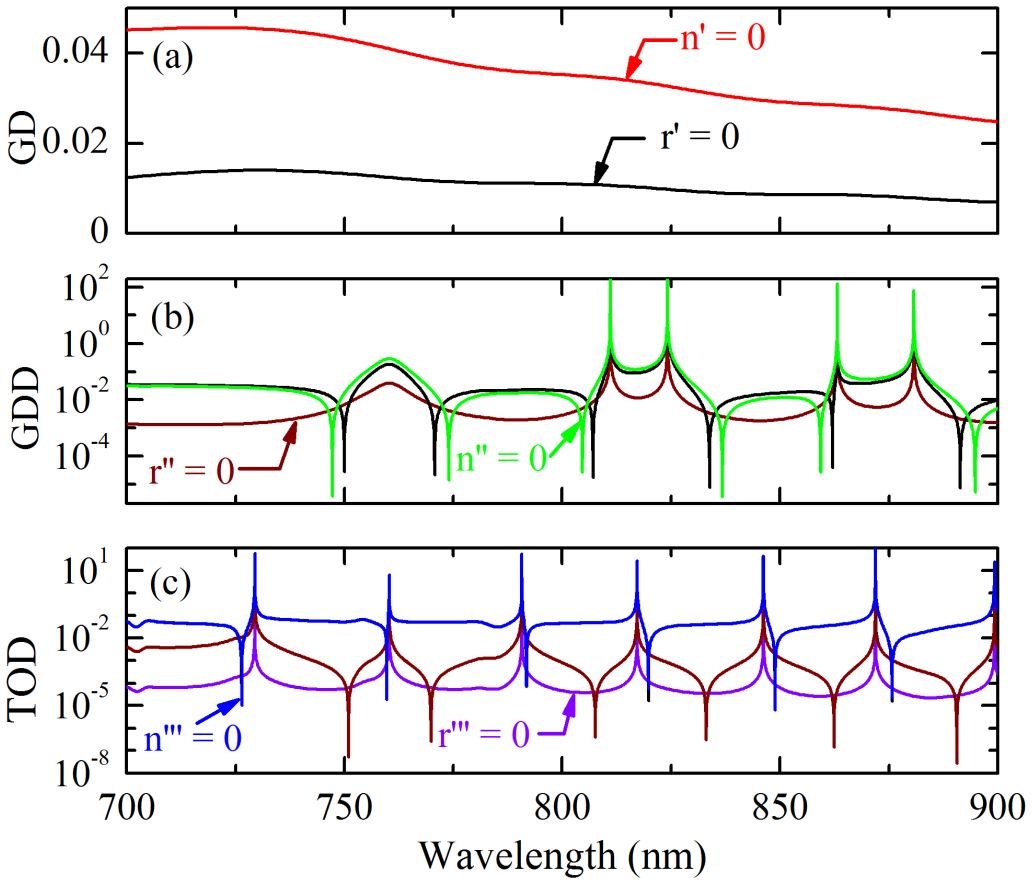

Figure 3.4: Relative errors calculated for the: (a) first, (b) second, and (c) third phase derivatives of the SiO_2-Nb_2O_5 chirped mirror structure of [26]. For each phase derivative, the impact on the relative error of assumptions made concerning the wavelength dependence of n or r for an individual material layer is shown by separate curves. All calculations assume normal incidence. In part (c), the $n''' = 0$ curve entirely covers the $r' = 0$ curve.

3.3.3 Posedness of Problem

The approach underlying the transfer matrix method is first solving Maxwell's equations in each individual layer of a mirror and subsequently joining the solutions through the boundary conditions. If a particular layer is lossy, or the optical field is evanescent in that layer, then the field will comprise two waves, one of which decays exponentially and another that grows exponentially. Error introduced to the calculated forward/backward ratio of the propagating waves will also grow exponentially as the forward wave propagates through the layer (and subsequent layers). In contrast, the Markov-Airy formalism

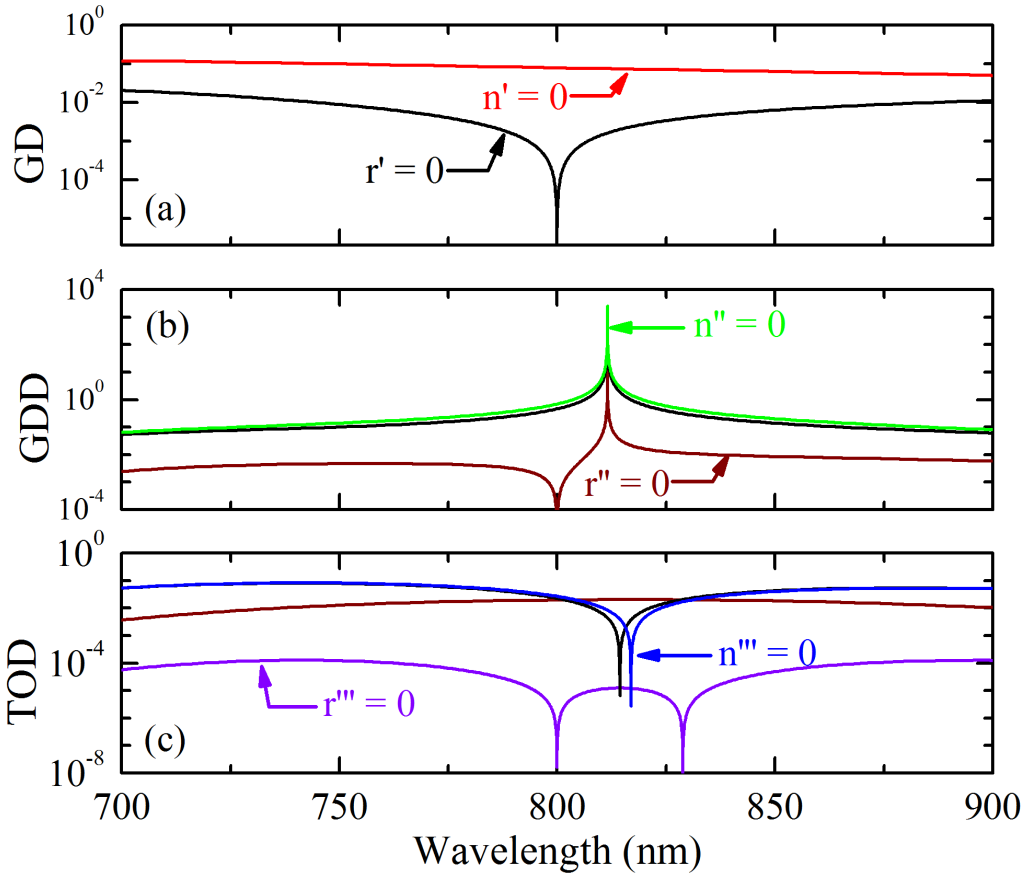

Figure 3.5: Calculations, similar to those of Figure 3.4 for the relative errors associated with phase derivative calculations (those for the first, second, and third derivatives are shown in panels (a), (b), and (c) respectively) but for a mirror comprising only 60 SiO_2-TiO_2 layer pairs. The center wavelength was chosen to be 800 nm and the angle of incidence in this case is assumed to be 45 degrees.

accounts for reflection and transmission at each layer interface. Since evanescent and lossy optical fields decay exponentially in each layer, the associated error will also not grow as the optical field traverses the mirror structure.

The posedness of a problem can be quantified by the condition number [28] which represents an upper bound on the multiplication of error when a linear system is solved. Furthermore, the log (base 10) of the condition number is interpreted as an upper bound on the number of decimal digits lost in performing computations on the system. In an effort to assess the conditioning of both the transfer matrix and the Markov-Airy formalisms, several additional mirror structures were examined. Figure 3.6 compares

the number of digits lost in numerical simulations of a mirror comprising 200 SiO_2/Ta_2O_5 layer pairs deposited onto a glass (n=1.5) substrate. The calculations assume SiO_2 and Ta_2O_5 layer thicknesses of 137.9 nm and 88.9 nm, respectively, and the angle of incidence is varied from zero to 90 degrees. Results are shown for both TE and TM polarization of the incident optical wave, and it is apparent that both the transfer matrix and Markov-Airy methods are well-posed when the angle of incidence is $< 70°$. At higher angles, however, the Markov formalism continues to be accurate but the condition number of the transfer matrix method is unbounded as the angle of incidence approaches 90°.

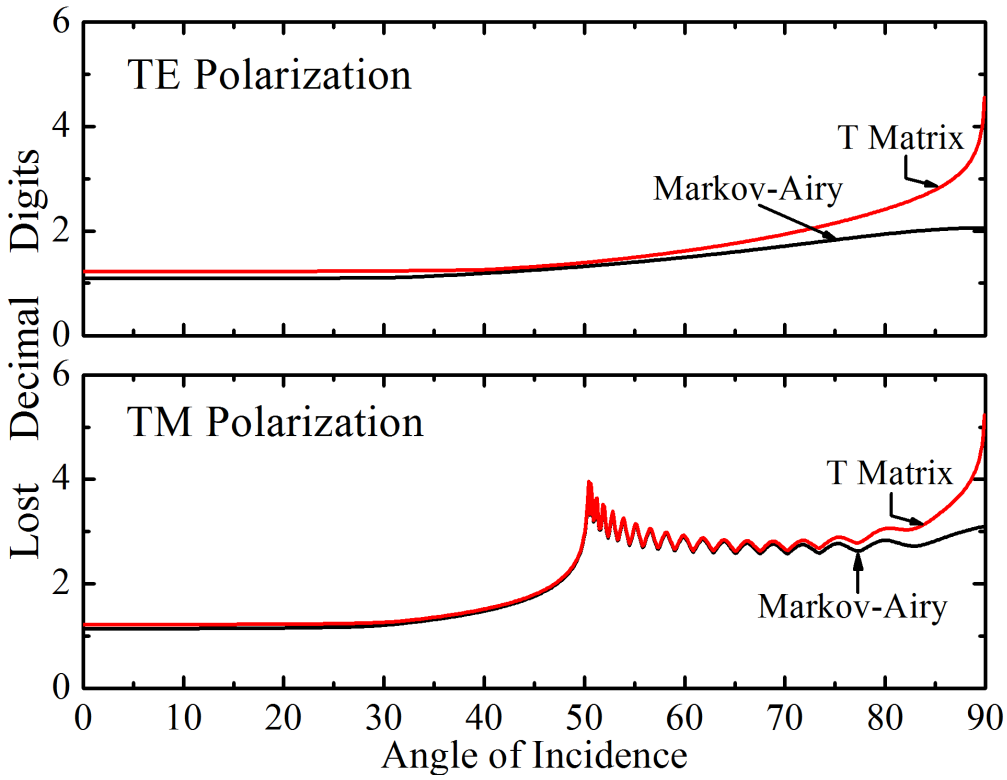

Figure 3.6: Comparison of the digits lost in computing the optical characteristics of ultra-broadband mirrors with the transmission matrix approach (shown in red), as opposed to the Markov-Airy formalism described here (blue). Shown as a function of the angle of incidence, these calculations assume the mirror comprises 200 SiO_2/Ta_2O_5 layer pairs deposited onto a glass (n=1.5) substrate. Incoming light is incident normal to the multilayer mirror from air: (Top) TE polarization of the incident optical wave, and (Bottom) TM polarization.

3.4 Scattering in Cylindrical and Spherical Coordinates

3.4.1 Introduction

The scattering formalism may also be used to calculate scattering from structures of cylindrical and spherical symmetry. There are several differences in scattering from a cylindrical or spherical object as compared to a planar structure. In planar structures, waves do not scatter among TE and TM polarizations, but in structures of cylindrical symmetry, those consisting entirely of isotropic materials, they do. Next, because structures of cylindrical or spherical symmetry contain an origin (the $\hat{z}$-axis and origin, respectively), there are no transmitted fields when calculating the scattering from a wave propagating toward the origin. Finally, the 'incident' wave may comprise components propagating in both the $+\hat{r}$ and $-\hat{r}$ directions. In calculating the scattering from plane waves, the plane waves are expanded as summations of Bessel functions of the first type, which can in turn be written as the sum of Hankel functions of the first and second kind.

In structures of both spherical and cylindrical symmetry, there are two types of scattering to consider. The first type is scattering of a wave incident on a shell from the inside. A sample structure and state diagram for this type of scattering, shown in Figure 3.7a, looks very much like the planewave version, with curved interfaces. The capital letters and thick lines indicate that the reflection and transmission coefficients are matrices and that the states are two-element vectors. The corresponding transition matrix for the structure is:

$$T_{out} = \begin{pmatrix} 0 & R_{12} & T_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_{23} & T_{32} & 0 \\ 0 & T_{12} & R_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & T_{23} & R_{32} & \end{pmatrix}. \tag{3.46}$$

The second type of scattering occurs when waves are incident from the outside onto a structure which contains the origin. An example structure and state diagram are shown in Figure 3.7b. The transition matrix for this structure

is:

$$T_{in} = \begin{pmatrix} 0 & R_{23} & T_{32} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_{23} \\ 0 & R_{23} & R_{32} & 0 \end{pmatrix}. \tag{3.47}$$

The fact that no transmitted wave exists does not imply that $|r| = 1$. If the structure contains any lossy materials, the scattering will be less.

To solve for the reflection coefficients from these structures, the T matrices must be converted into Q matrices, as was done in the planar case. Once again, the technique of adding a zero-thickness layer to the structure, shown in Figure 3.2 for planar coordinates, should be used.

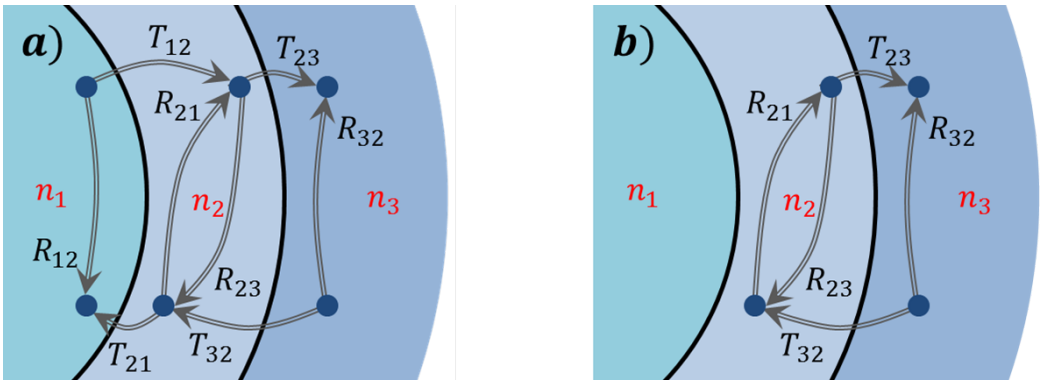

Figure 3.7: (a) State diagram for scattering of an outward-propagating wave. (b) State diagram for scattering of an inward-propagating wave. The double lines indicate the two propagating (and intermixing) polarizations.

In spherical coordinates, as well as when m=0 or k_z=0 in cylindrical coordinates, polarizations do not mix at interfaces. In these cases, the state diagrams and transition matrices may be separated into two independent systems for TE and TM polarizations. The state diagrams will appear identical to those in Figure 3.7, and the transition matrices will be the same as those in Equations 3.46 and 3.47, but with the vectors and matrices replaced by scalar quantities. The appropriate excitation vectors for the outward and inward waves are

$$E_{0,out} = (0, 1, 0, |\ldots 0, 0)^T \tag{3.48}$$

$$E_{0,in} = (0, 0|\ldots|0, 1, 0)^T, \tag{3.49}$$

respectively. When Equation 3.30 is solved, the reflection coefficient is $E_2^{(0)}$

for the outwardly traveling wave and $E^{(0)}_{end-1}$ for the inwardly traveling wave, using the same convention to index $E^{(0)}$ as in Equation 3.10.

An approach similar to that used in Equation 2.83 is adopted to calculate the scattering for $m \neq 0$ in cylindrical coordinates. The structure is excited separately with TE and TM waves to determine the coefficients r_{ee}, r_{em}, r_{me}, and r_{mm}. The appropriate excitation vectors are:

$$E^{TE}_{0,out} = (0,0|1,0|0,0|\ldots|0,0)^T \tag{3.50}$$

$$E^{TM}_{0,out} = (0,0|0,1|0,0|\ldots|0,0)^T \tag{3.51}$$

$$E^{TE}_{0,in} = (0,0|\ldots|0,0|1,0|0,0)^T \tag{3.52}$$

$$E^{TM}_{0,in} = (0,0|\ldots|0,0|0,1|0,0)^T. \tag{3.53}$$

For example, when $Q_{out}E^{(0)} = E^{TE}_{0,out}$ is solved, r_{ee} and r_{em} correspond to $E^{(0)}_3$ and $E^{(0)}_4$, respectively.

At this point, we would like to step back and discuss the meaning of the reflection and transmission coefficients. Consider the scattering of an inwardly traveling spherical wave incident on a structure of spherical symmetry. When we say the scattering coefficient is r_s, we mean that if the incident electric vector potential has the form $F_r = \hat{H}^{(2)}_\ell(kr)P^m_\ell(\cos\theta)e^{jm\phi}$, then the reflected electric vector potential is $F_r = r_s\hat{H}^{(1)}_\ell(kr)P^m_\ell(\cos\theta)e^{jm\phi}$.

The theory of a plane wave scattering from a sphere is known as Mie theory [29, 30]. A similar theory exists for scattering by cylindrical objects. In Mie theory, the incident plane wave is written as an infinite sum of Riccati-Bessel functions of the first kind ($\hat{J}_\ell$), and the scattered field is written as the sum of Riccati-Hankel functions of the first[2] kind $\hat{H}^{(1)}_\ell$.

We seek to use the Markov theory of this section to calculate Mie scattering. To do so, let the reflection coefficient from $\hat{H}^{(2)}_\ell$ to $\hat{H}^{(1)}_\ell$ be r_s and the reflection from $\hat{J}_\ell$ to $\hat{H}^{(1)}_\ell$ be r_{mie}. Consider an imaginary boundary, positioned just outside the outermost boundary of the structure. Inside the boundary, the radial part of the electric vector potential takes the form $\hat{H}^{(2)}_\ell(kr)+r_s\hat{H}^{(1)}_\ell$. Outside the boundary, it takes the form $\hat{J}_\ell(kr) + r_{mie}\hat{H}^{(1)}_\ell$. Using Table 2.6

[2]Sources that use $e^{j\omega t}$ time dependence will use $\hat{H}^{(2)}_\ell$ for the scattered field.

to write the continuity equations for E_ϕ and H_θ, we find:

$$\hat{J}_\ell(kr) + r_{mie}\hat{H}^{(1)}(kr) = t\left(\hat{H}^{(2)}(kr) + r\hat{H}^{(1)}(kr)\right) \tag{3.54}$$

$$\hat{J}'_\ell(kr) + r_{mie}\hat{H}^{(1)\prime}(kr) = t\left(\hat{H}^{(2)\prime}(kr) + r\hat{H}^{(1)\prime}(kr)\right). \tag{3.55}$$

When Equations 3.54 and 3.55 are solved for r_{mie}, the result, when simplified by means of the Wronskian relations[3] for Riccati-Bessel functions, is

$$r_{mie} = \frac{1}{2}(r_s - 1). \tag{3.56}$$

Equation 3.56 also applies for TM reflection coefficients. Furthermore, it may be used to relate r_s to r_{mie} for spheres coated with layers of shells.

There is one last consideration when working in cylindrical and spherical coordinates. When their argument is complex, $H_m^{(1)}$ and $\hat{H}_\ell^{(1)}$ become very small and $H_m^{(2)}$ and $\hat{H}_\ell^{(2)}$ become very large. When calculating reflection and transmission coefficients from inside a shell where the wave is evanescent, the resulting reflection and transmission coefficients may be very large or small. In such cases, not all numerical linear algebra packages will solve Equation 3.30 correctly. To remedy this problem, the fields may be renormalized. As an example, consider a thin shell in cylindrical coordinates in which the field is evanescent which extends from radius a to radius b. The outward propagating part of the field is renormalized to $H_m^{(1)}(ka)/H_m^{(1)}(ka) = 1$ at the $r = a$ boundary and as $H_m^{(1)}(kb)/H_m^{(1)}(ka)$ at the $r = b$ boundary. Similarly, for the inward propagating part of the field, the field is renormalized as $H_m^{(2)}(ka)/H_m^{(2)}(kb)$ at the $r = a$ boundary and $H_m^{(2)}(kb)/H_m^{(2)}(kb) = 1$ at the $r = b$ boundary. To calculate the fields inside the structure, the rescaling must be reversed.

3.4.2 Implementation and Validation

To demonstrate the correctness of the algorithm, code to calculate scattering from structures possessing cylindrical or spherical symmetry was implemented in MATLAB. The calculated Mie reflection coefficients were com-

[3] $W\left\{\hat{J}_\ell(z), \hat{Y}_\ell(z)\right\}=1$, $W\left\{\hat{J}_\ell(z), \hat{H}_\ell^{(1)}(z)\right\}=j$, $W\left\{\hat{J}_\ell(z), \hat{H}_\ell^{(2)}(z)\right\}=-j$, $W\left\{\hat{H}_\ell^{(1)}(z), \hat{H}_\ell^{(2)}(z)\right\}=-j2$

pared with the values from a formula valid for a single sphere [30]. A typical result is shown in Figure 3.8, which presents the calculated Mie reflection coefficients for TE waves. A similar graph exists for TM waves. Note that the error is near the limit of the machine's accuracy.

The code was more thoroughly validated than the one test case (an uncoated sphere) would suggest. Zero thickness layers were added both inside and outside the spheres. Mathematically, these layers should not affect the result, but by adding them, they test the code's ability to correctly create transition matrices. Furthermore, the validation of the code for calculating the propagation constant in cylindrical waveguides and the frequency of whispering gallery modes (to be discussed in Chapter 4) in spheres serves to further validate the code.

3.5 Rigorous Coupled Wave Analysis

Scattering from arbitrary structures may be calculated with rigorous coupled wave analysis (RCWA), which may be formulated in terms of the Markov-Airy method. The theory of RCWA has been described in detail elsewhere [6, 31], and the procedure is summarized here: 1) The structure is divided into many thin layers, such that variations in the cross section in any layer are much smaller than the wavelength. 2) The eigenmodes of propagation are then calculated in each layer, commonly with planewave expansions. 3) At the boundaries, scattering coefficients among the eigenmodes of the two layers are computed. 4) The scattering coefficients of the individual layers are combined to determine the scattering of the entire structure, as described in Reference [31].

Step 4 above may be replaced with the Markov-Airy method. A state diagram may be drawn and is shown in Figure 3.9. The transition matrix

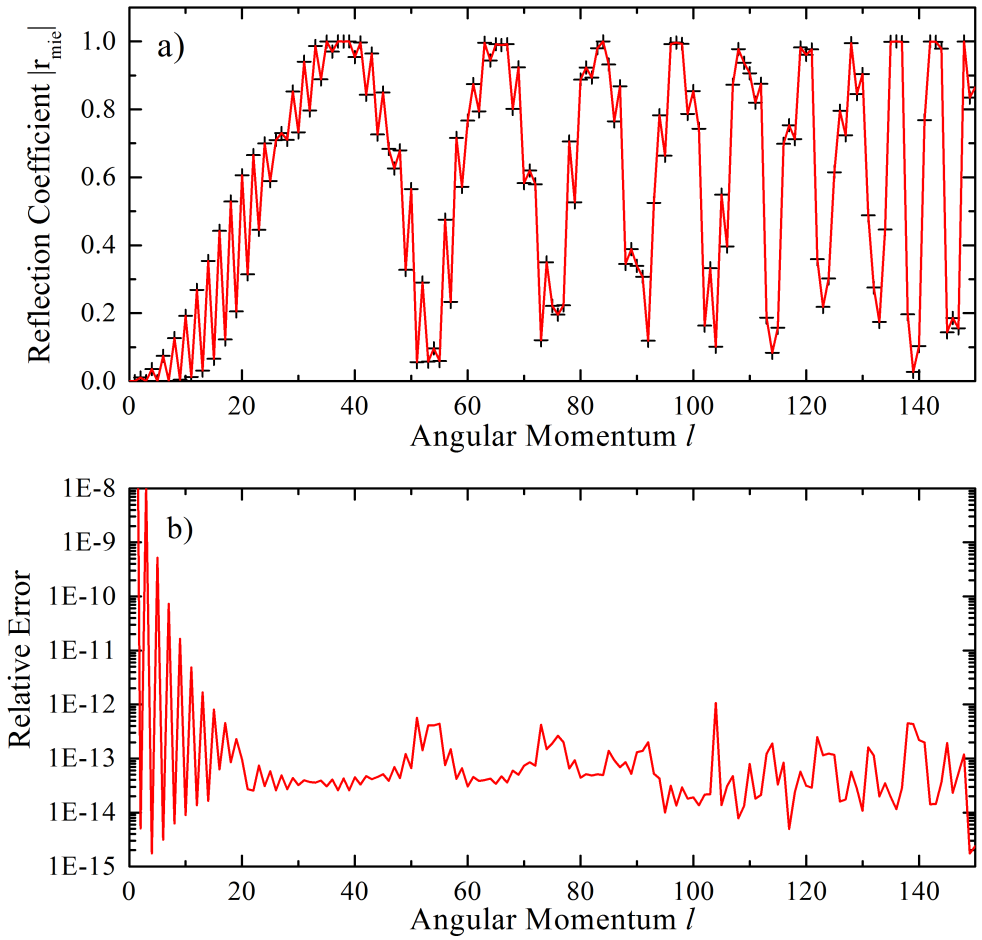

Figure 3.8: (a) Mie reflection coefficients for 400 nm light incident upon a 10 μm sphere with refractive index n=1.5. The red curve is the value while the '+' marks indicate the values computed by the Markov algorithm. (b) The relative error between the two plots. The large values near $\ell = 0$ are the result of the reflection coefficient being nearly zero.

for the structure is:

$$T = \begin{pmatrix} I & R_{12} & T_{21} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & R_{23}e^{j\Phi} & T_{32}e^{j\Phi} & 0 \\ 0 & T_{12}e^{j\Phi} & R_{21}e^{j\Phi} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & T_{23} & R_{32} & I \end{pmatrix}. \tag{3.57}$$

As compared to Equation 3.2, the most obvious difference is that all elements

are matrices instead of scalars, which is indicated by the use of capital letters. R_{mp} and T_{mp} describe the scattering among the various eigenmodes of two layers at the boundary between them. The phase factor, Φ , is a diagonal matrix, comprising the propagation constants of all of the modes multiplied by the thickness of the layer. The Q matrix is defined similarly to Equation 3.3, representing the transitions between only transient states. The excitation vector, E_0, is computed by expanding the exciting fields in the basis of the propagating eigenmodes. The field in the entire structure is found by solving

$$(I - Q)E^{(0)} = E_0, \tag{3.58}$$

where $E^{(0)}$ is expressed in the eigenmode basis of all of the layers.

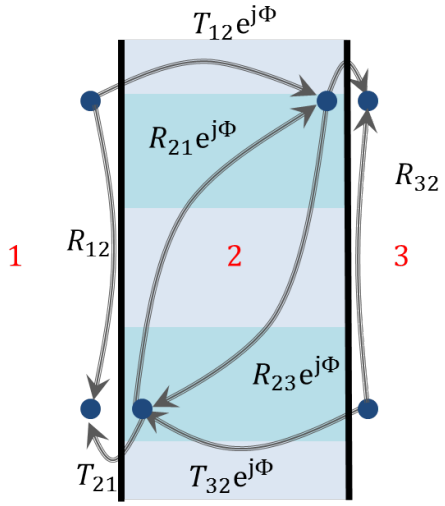

Figure 3.9: State diagram of RCWA scattering problem.

3.6 Conclusions

A new method for analyzing multilayer optical structures, such as chirped grating mirrors for ultrafast laser applications, has been presented. Based on the application of Markov chains to Airy's analysis, the formalism is exact, and the computational complexity of the technique scales linearly with both the number of mirror layers and the order of the highest phase derivative to be calculated. Exact expressions for the three lowest-order phase derivatives have been presented, and the impact of simplifying assumptions regarding

the frequency dependence of n and r have been discussed. Since femtosecond and attosecond laser technology is placing ever-increasing demands on the properties of optical components, the Markov-Airy formalism reported here is promising for the design of multilayer mirror structures offering improved dispersion characteristics.

An adaptation of the new method to cylindrical and spherical coordinates has also been discussed. As with planar theory, calculating scattering with the Markov-Airy method should be a better-posed problem than doing so with the conventional T-matrix approach.

CHAPTER 4

ANALYSIS OF WAVEGUIDES AND RESONATORS

In Chapter 3, it was shown how one might calculate the scattering of free-propagating light by a multilayer structure. In addition to scattering light, multilayer structures may also guide light parallel to the layer interfaces via reflection. These are known as waveguide modes, and may be lossy or lossless. In structures with cylindrical and spherical symmetry, the light may be waveguided back upon itself, creating a closed loop. These types of modes are known as whispering gallery modes. In this chapter, the theory of how to find waveguide and resonator modes will be discussed in rectangular, cylindrical and spherical coordinates.

4.1 Lossless Waveguides

Eigenmodes of lossless waveguides propagate without changing shape or amplitude. For a waveguide to be lossless, it must be constructed from materials which are not intrinsically lossy (do not absorb or scatter the light). Furthermore, such waveguides must possess at least one layer with a greater index of refraction than that of the surrounding medium. With such a layer, it may guide the light with total internal reflection.

At a fixed frequency, lossless waveguides possesses a finite number of discrete modes, each of which is associated with a propagation constant $k_z^{(m)}$. The propagation constant must be real because if it were complex, the field would be exponentially growing or decaying along the $\hat{z}$ direction, which is in violation of energy conservation. All mode constants must fall into the range

$$\frac{\omega n_{sub}}{c} < k_z^{(m)} < \frac{\omega n_{max}}{c}, \tag{4.1}$$

where n_{sub} and n_{max} are the refractive indices of the substrate and the layer having the highest index, respectively. If $k_x \leq \frac{\omega n_{sub}}{c}$, the field would not be

evanescent in the surrounding medium and if $k_z \geq \frac{\omega n_{max}}{c}$, the field could not propagate in any layer. The solutions will be indexed by m and referred to as k_x^m, with $m = 0$ corresponding to the largest propagation constant. The physical significance of the mode index m is that it represents the number of nodes in the solution field.

4.1.1 The Phase Method for Planar Waveguides

The structure of a planar waveguide is shown in Figure 4.1a. Layers of materials are assumed to be infinite in both the $\hat{y}$ and $\hat{z}$ planes. Because the index profile may be written as the product $n(x, y, z) = n_x(x)n_z(y)n_z(z)$, the technique of separation of variables may be used.[1] Without loss of generality, the wave may be assumed to propagate in the $\hat{z}$ direction. In this case, the field will take the form $E(x)e^{jk_z z}$. Layer thicknesses and indices of refraction are given and the problem is to find the propagation constants $k_z^{(m)}$ as well as the associated electric fields.

For any allowed mode of a planar waveguide, the round trip phase shift in the transverse direction must be a multiple of 2π. The reason for this conclusion is that the electric field must be single-valued. If this equation is applied to a layer indexed by q in the waveguide, the dispersion relation takes the form:

$$\angle R_a(k_z, \omega) + \angle R_b(k_z, \omega) + 2\phi_q(k_z, \omega) = 2\pi m, \tag{4.2}$$

where R_a is the reflection from all of the layers above layer q and including the phase of layer q, R_b is the reflection from all of the layers below layer q, and m is the mode index. It will be convenient to choose the layer q which has the highest index of refraction in the structure, to ensure the field is not evanescent.

A waveguide may be pictured as two mirrors stuck together. As soon as a wave is reflected out of one mirror structure, it immediately enters the other mirror. This point of view enables us to use the results of Chapter 3 to find the modes of the waveguide and their dispersion. To do so, an imaginary, zero-thickness layer with an index of refraction greater than or equal to that of the largest layer is inserted anywhere in the structure (it could even be

[1] See Section 2.2.1 for details

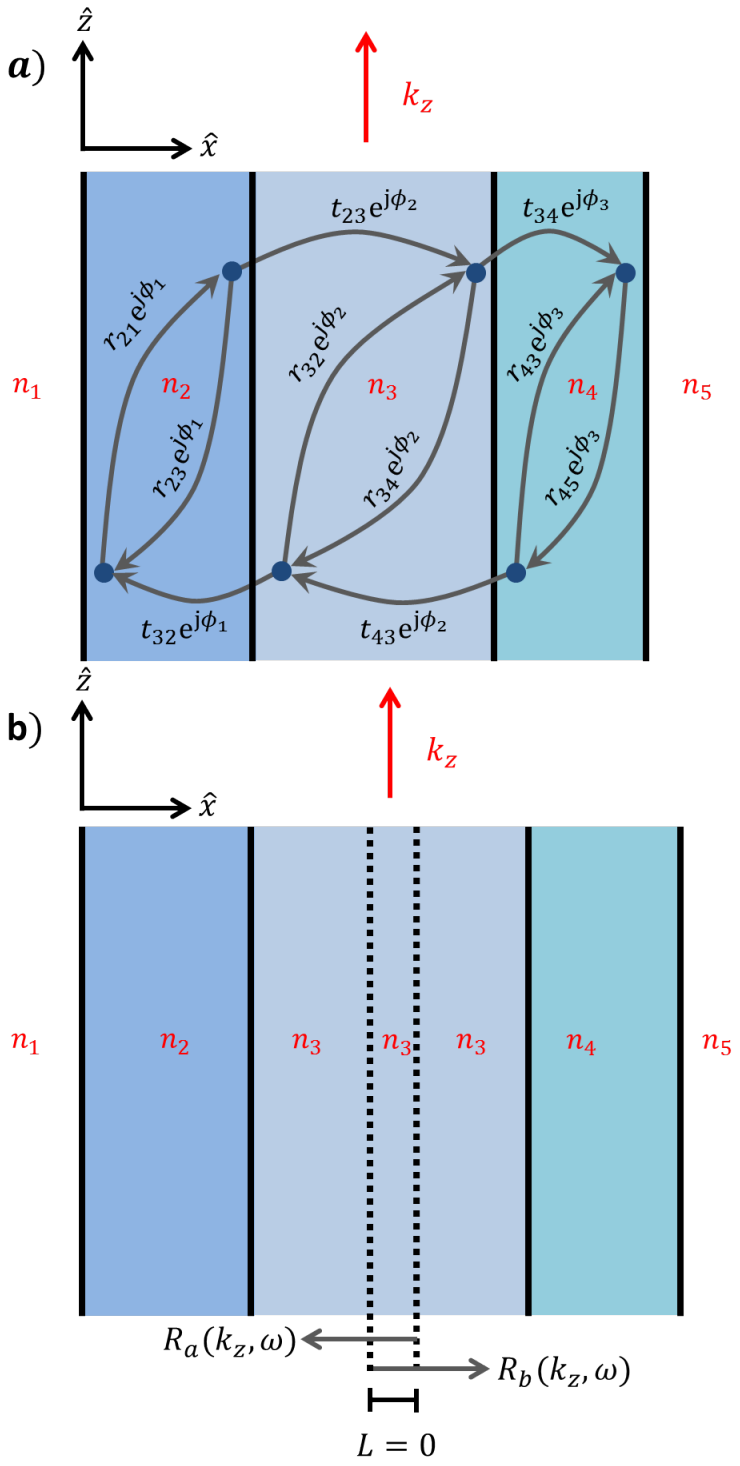

Figure 4.1: (a) Planar waveguide with a superimposed Markov state diagram. (b) The same waveguide showing an imaginary (indicated by the dashed lines), zero-thickness layer inserted into one of the layers.

outside the waveguide). This layer defines the boundary between the two mirrors. Because the layer is of zero thickness, it will not affect the field distribution and the phase term in Equation 4.2 is zero. The dispersion relationship becomes

$$\angle R_a(k_z, \omega) + \angle R_b(k_z, \omega) = 2\pi m. \tag{4.3}$$

Figure 4.1b illustrates the use of Equation 4.3.

Total internal reflection occurs for both TE and TM waves. In planar structures, polarization is not mixed by reflections at the boundaries. TE and TM waveguide propagation constants may be located by solving Equation 4.3 with TE and TM reflection coefficients.

The need for the field to remain single-valued will remain even in lossy waveguides and other coordinate systems. The technique of inserting an imaginary, high-index layer inside the structure will be used throughout the remainder of this chapter. Not only does this simplify the equations, but it also makes the equations share the same form among various coordinate systems.

4.1.2 Dispersion in Planar Waveguides

In addition to finding the propagation constants themselves, it is interesting to examine how the propagation constants vary as a function of frequency (group delay, group delay dispersion, and third order dispersion). These quantities characterize how a waveguide will delay and temporally distort light which propagates through it.

Consider the dispersion relations. For the phase method, the dispersion relationship is:

$$f(k_z, \omega) \equiv \angle R_a(k_z, \omega) + \angle R_b(k_z, \omega) = 2\pi m. \tag{4.4}$$

Because the propagation constant is constrained to be real, the propagation constants may be found for by sampling f at fixed ω_0 in the k_z range specified by Equation 4.1 and looking for intervals where the real and imaginary parts of f change sign. While this method is simple, it has the disadvantage that two closely spaced modes may be missed. The more general method to be

discussed in Section 4.2 may be used. It deterministically finds all roots, but at the cost of greater complexity and computational effort.

It is not generally possible to solve for the propagation constants as a function of frequency. To determine the dispersion, implicit partial differentiation is adopted. For the m^{th} mode, the group delay per unit length is:

$$\frac{dk_z}{d\omega} = -\frac{f_\omega(k_z^{(m)}, \omega_0)}{f_{k_z}(k_z^{(m)}, \omega_0)}. \tag{4.5}$$

To find the group delay dispersion and the third order dispersion, the total derivatives of Equation 4.5 must be found. The results are:

$$\frac{d^2k_z}{d\omega^2} = \frac{2f_{k_z}f_\omega f_{k_z\omega} - f_{k_zk_z}f_\omega^2 - f_{k_z}^2 f_{\omega\omega}}{f_{k_z}^3} \tag{4.6}$$

$$\begin{aligned}\frac{d^3k_z}{d\omega^3} = &-f_{k_z}^{-5}\left[3f_\omega^3 f_{k_zk_z}^2 + f_{k_z}^4 f_{\omega\omega\omega} + 3f_{k_z}^2 f_\omega^2 f_{k_zk_z\omega} - f_{k_z}f_\omega^3 f_{k_zk_zk_z} + 6f_{k_z}^2 f_\omega f_{k_z\omega}^2\right.\\ &\left.-3f_{k_z}^3 f_\omega f_{k_z\omega\omega} - 3f_{k_z}^3 f_{k_z\omega}f_{\omega\omega} - 9f_{k_z}f_\omega^2 f_{k_zk_z}f_{k_z\omega} + 3f_{k_z}^2 f_\omega f_{k_zk_z}f_{\omega\omega}\right].\end{aligned} \tag{4.7}$$

For compactness, the arguments of the derivatives of f have been suppressed. All expressions should still be evaluated at $(k_z^{(m)}, \omega_0)$.

In order to use Equations 4.5-4.7, the phase and phase derivatives for reflection of the two 'mirrors' which comprise the waveguide must be calculated, including partial phase derivatives. Following the reasoning in Sections 3.2.2 and 3.2.3, various partial derivatives are taken of Equation 3.12. The terms are then rearranged to provide systems of equations, which, when solved, yield the complex derivatives of the reflectivity vector. The derivatives with respect to frequency are:

$$(I-Q)E_\omega^{(1)} = Q_\omega^{(1)}E^{(0)} \tag{4.8}$$

$$(I-Q)E_{\omega\omega}^{(2)} = Q_{\omega\omega}^{(2)}E^{(0)} + 2Q_\omega^{(1)}E_\omega^{(1)} \tag{4.9}$$

$$(I-Q)E_{\omega\omega\omega}^{(3)} = Q_{\omega\omega\omega}^{(3)}E^{(0)} + 3Q_{\omega\omega}^{(2)}E_\omega^{(1)} + 3Q_\omega^{(1)}E_{\omega\omega}^{(2)}. \tag{4.10}$$

Analogous to those for frequency, the partial derivatives with respect to k_z

may be found by solving:

$$(I-Q)E^{(1)}_{k_z} = Q^{(1)}_{k_z}E^{(0)} \tag{4.11}$$
$$(I-Q)E^{(2)}_{k_zk_z} = Q^{(2)}_{k_zk_z}E^{(0)} + 2Q^{(1)}_{k_z}E^{(1)}_{k_z} \tag{4.12}$$
$$(I-Q)E^{(3)}_{k_zk_zk_z} = Q^{(3)}_{k_zk_zk_z}E^{(0)} + 3Q^{(2)}_{k_zk_z}E^{(1)}_{k_z} + 3Q^{(1)}_{k_z}E^{(2)}_{k_zk_z}. \tag{4.13}$$

Finally, the systems of equations for finding partial derivatives are:

$$(I-Q)E^{(2)}_{k_z\omega} = Q^{(2)}_{k_z\omega}E^{(0)} + Q^{(1)}_{k_z}E^{(1)}_{\omega} + Q^{(1)}_{\omega}E^{(1)}_{k_z} \tag{4.14}$$
$$(I-Q)E^{(3)}_{k_zk_z\omega} = Q^{(3)}_{k_zk_z\omega}E^{(0)} + 2Q^{(2)}_{k_z\omega}E^{(1)}_{k_z} + Q^{(2)}_{k_zk_z}E^{(1)}_{\omega} + 2Q^{(1)}_{k_z}E^{(2)}_{k_z\omega} + Q^{(1)}_{\omega}E^{(2)}_{k_zk_z} \tag{4.15}$$
$$(I-Q)E^{(3)}_{k_z\omega\omega} = Q^{(3)}_{k_z\omega\omega}E^{(0)} + 2Q^{(2)}_{k_z\omega}E^{(1)}_{\omega} + Q^{(2)}_{\omega\omega}E^{(1)}_{k_z} + 2Q^{(1)}_{\omega}E^{(2)}_{k_z\omega} + Q^{(1)}_{k_z}E^{(2)}_{\omega\omega}. \tag{4.16}$$

Note that $E^{(2)}_{k_z\omega} = E^{(2)}_{\omega k_z}$ and $Q^{(2)}_{k_z\omega} = Q^{(2)}_{\omega k_z}$. Although Equations 4.8-4.10 appear similar to Equations 3.13-3.15, they are different in that in the former equation, $Q^{(1)}_{\omega}$ represents a partial derivative, whereas in Equations 3.13-3.15, $Q^{(1)}$ represents a total derivative.

The partial derivatives of the matrices are given by the matrix formed by taking the partial derivatives of the elements. Again, elements of the matrix take the form of Equation 3.16. The derivatives of these elements are:

$$\Delta Q^{(2)}_{k_z\omega} = e^{j\phi}\left[r_{k_z\omega} + j\phi_{k_z}r_{\omega} + j\phi_{\omega}r_{k_z} + \left(j\phi_{k_z\omega} - \phi_{k_z}\phi_{\omega}\right)r\right] \tag{4.17}$$
$$\begin{aligned}\Delta Q^{(3)}_{k_zk_z\omega} = e^{j\phi}\Big[&r_{k_zk_z\omega} + j2\phi_{k_z}r_{k_z\omega} + j\phi_{\omega}r_{k_zk_z} + \left(j\phi_{k_zk_z} - \phi^2_{k_z}\right)r_{\omega} \\ &+2\left(j\phi_{k_z\omega} - \phi_{k_z}\phi_{\omega}\right)r_{k_z} + \left(j\phi_{k_zk_z\omega} - 2\phi_{k_z\omega}\phi_{k_z} - \phi_{k_zk_z}\phi_{\omega} - j\phi^2_{k_z}\phi_{\omega}\right)r\Big]\end{aligned} \tag{4.18}$$
$$\begin{aligned}\Delta Q^{(3)}_{k_z\omega\omega} = e^{j\phi}\Big[&r_{k_z\omega\omega} + j2\phi_{\omega}r_{k_z\omega} + j\phi_{k_z}r_{\omega\omega} + \left(j\phi_{\omega\omega} - \phi^2_{\omega}\right)r_{k_z} \\ &+2\left(j\phi_{k_z\omega} - \phi_{k_z}\phi_{\omega}\right)r_{\omega} + \left(j\phi_{k_z\omega\omega} - 2\phi_{k_z\omega}\phi_{\omega} - \phi_{k_z}\phi_{\omega\omega} - j\phi_{k_z}\phi^2_{\omega}\right)r\Big],\end{aligned} \tag{4.19}$$

where terms have been grouped by r derivatives.

To find the partial derivatives of the reflection and transmission coefficients, a strategy is followed that is similar to the one used in Equations 3.39-3.42. The reflection and transmission coefficients may be represented by the quotient of two functions, with f as the numerator and g as the

denominator. The partial derivative with respect to k_z and ω is given by:

$$r_{k_z\omega} = \frac{f_{k_z\omega}g^2 - fgg_{k_z\omega} - f_{k_z}gg_\omega - f_\omega gg_{k_z} + 2fg_{k_z}g_\omega}{g^3}. \tag{4.20}$$

The expressions for the third order derivatives ($k_zk_z\omega$ and $k_z\omega\omega$) are too large to be presented here.

The derivatives of the phase with respect to frequency are identical to the formulae presented in Equations 3.43-3.45 with $k'_{z0} = 0$. The reason k'_{z0} is zero is because, in the waveguide problem, k_z and ω are treated as independent variables. When simulating reflection at a fixed angle of incidence, as was done in Chapter 3, k_{z0} was a function of ω. The partial derivatives with respect to k_z are:

$$\gamma_m = \phi_m = (k_m^2 - k_z^2)^{1/2} \tag{4.21}$$

$$\gamma_{k_z,m} = \frac{d\phi_m}{dk_z} = L_m\frac{-k_z}{(k_m^2 - k_z^2)^{1/2}} \tag{4.22}$$

$$\gamma_{k_z,m} = \frac{d^2\phi_m}{dk_z^2} = L_m\frac{-k_m^2}{(k_m^2 - k_z^2)^{3/2}}. \tag{4.23}$$

Next, the partial derivatives of the phase with respect to k_z and ω are

$$\gamma_{k_z\omega,m} = \frac{d^2\phi_m}{\partial k_z\partial\omega} = L_m\frac{k_zk_mk'_m}{(k_m^2 - k_z^2)^{3/2}} \tag{4.24}$$

$$\gamma_{k_zk_z\omega,m} = \frac{d^3\phi_m}{\partial k_z^2\partial\omega} = L_m\frac{k_mk'_m\,(k_m^2 + 2k_z^2)}{(k_m^2 - k_z^2)^{5/2}} \tag{4.25}$$

$$\gamma_{k_z\omega\omega,m} = \frac{d^3\phi_m}{\partial k_z\partial\omega^2} = L_m\frac{k_z(k_m^3k''_m - k_z^2k'^2_m - 2k_m^2k'^2_m - k_z^2k_mk''_m)}{(k_m^2 - k_z^2)^{5/2}}. \tag{4.26}$$

Once the complex derivatives of the reflection coefficients are determined from Equations 4.8-4.16, they must be converted to phase derivatives for use in Equations 4.5-4.7. Appendix B discusses the method for calculating

partial derivatives of phase. The results are:

$$r_{k_x\omega} = \bar{R}^{(2)}_{k_z\omega,\hat{r}} + r\theta_{k_z}\theta_\omega \tag{4.27}$$

$$\theta_{k_z\omega} = \frac{1}{r}\left(\bar{R}^{(2)}_{k_z\omega,\hat{\theta}} - r_\omega\theta_{k_z} - r_{k_z}\theta_\omega\right) \tag{4.28}$$

$$\theta_{k_z\omega\omega} = \frac{1}{r}\left(\bar{R}^{(3)}_{k_z\omega\omega,\hat{\theta}} - 2r_{k_z\omega}\theta_\omega - r_{\omega\omega}\theta_{k_z} - r_{k_z}\theta_{\omega\omega} - 2r_\omega\theta_{k_z\omega} + r\theta_{k_z}\theta^2_\omega\right) \tag{4.29}$$

$$\theta_{k_zk_z\omega} = \frac{1}{r}\left(\bar{R}^{(3)}_{k_zk_z\omega,\hat{\theta}} - 2r_{k_z\omega}\theta_{k_z} - r_{k_zk_z}\theta_\omega - r_\omega\theta_{k_zk_z} - 2r_{k_z}\theta_{k_z\omega} + r\theta^2_{k_z}\theta_\omega\right). \tag{4.30}$$

4.1.3 Determinant Method for Planar Waveguides

First, the problem of waveguides will discussed with an extension to the Markov analogy developed in the previous chapter. If a scattering multilayer structure corresponds to a Markov chain with absorbing states, a waveguide is a non-absorbing Markov chain. The problem of finding the waveguide modes then reduces to finding the steady state solution of the transition matrix. It is known that, since the field must not change upon reflection, the eigenvalue of the transition matrix must be one. A diagram of a planar waveguide is shown in Figure 4.1a with a state diagram superimposed on top of it. As in Chapter 3, a transition matrix for the system may be formed. It is given by:

$$Q(k_z,\omega) = \begin{pmatrix} 0 & r_{23}e^{j\phi_1} & t_{32}e^{j\phi_1} & 0 & 0 & 0 \\ r_{21}e^{j\phi_1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & r_{34}e^{j\phi_2} & t_{43}e^{j\phi_2} & 0 \\ 0 & t_{23}e^{j\phi_2} & r_{32}e^{j\phi_2} & 0 & 0 & 0 \\ 0 & 0 & 0 & & 0 & r_{45}e^{j\phi_3} \\ 0 & 0 & 0 & t_{34}e^{j\phi_3} & t_{43}e^{j\phi_3} & 0 \end{pmatrix}. \tag{4.31}$$

Here $|r_{21}| = |r_{34}| = 1$, because for a lossless waveguide mode, the field must be totally internally reflected inside the structure. Note that the matrix Q is a function of both k_z and ω, because both the phase and reflection coefficients depend on those parameters. Because 1 must be an eigenvalue of a transition

matrix for a guided mode,

$$f(k_z, \omega) \equiv \det(I - Q(k_z, \omega)) = 0 \tag{4.32}$$

for any mode. The above equation defines the dispersion relationship. The propagation constants may be found by locating the roots of f, as was done in Section 4.1.1. However, finding the determinant of large matrices is not a numerically stable operation, making this technique unsuitable for many-layered structures.

To calculate the derivatives of the propagation constant using Equations 4.5-4.7, it is necessary to find the derivative of a determinant. This may be accomplished with Jacobi's identity:

$$\frac{d}{dx}|A(x)| = tr\left(adj(A(x))\frac{dA(x)}{dx}\right), \tag{4.33}$$

where tr represents the trace, adj is the operation of taking the adjunct (the transpose of the cofactor matrix), and $dA(x)/dx$ is the matrix whose elements are the derivatives of those in $A(x)$. The cofactor matrix is relatively computationally expensive to implement, i.e. $\mathcal{O}(n^3)$.

Compared to the phase method for finding the modes of waveguides and calculating the dispersion, the determinant method is slower and not numerically stable for many-layered structures. This method has been discussed here because the reasoning outlined above will allow us to write the dispersion relations for cylindrical waveguides and spherical whispering gallery modes, which is not otherwise obvious because TE and TM modes intermix.

4.1.4 Phase Method for Cylindrical Waveguides

The theory of waveguides with cylindrical symmetry is broadly similar to that of planar waveguides. A fictitious, high-index layer is inserted into the structure. From inside this layer, the reflection coefficients toward and away from the $\hat{z}$ axis are calculated.

When m =0 in cylindrical coordinates, no scattering occurs between TE^z and TM^z modes at the interfaces between materials and, as a result, the TE and TM modes are unmixed and may be located separately using Equation 4.3. The derivatives of the scattering coefficients (Equations 2.88-2.89 and

2.92-2.93) may be found with the same strategy presented in Equations 3.39-3.42 and 4.20.

For $m \neq 0$ modes, a solution is complicated by scattering between TE^z and TM^z modes. The reflection coefficients toward and away from the $\hat{z}$-axis are calculated and used to construct the non-absorbing transition matrix:

$$T(k_z, \omega) = \begin{pmatrix} 0 & R_A \\ R_B & 0 \end{pmatrix}, \tag{4.34}$$

where R_A and R_B are 2x2 matrices. When $\det(I - T)$ is expanded, the resulting eigenequation is:

$$\begin{aligned} 0 = f(k_z, \omega) = {} & r_{ee,a} r_{ee,b} + r_{mm,a} r_{mm,b} + r_{em,a} r_{me,b} + r_{me,a} r_{em,b} \\ & - (r_{ee,a} r_{mm,a} - r_{em,a} r_{me,a})(r_{ee,b} r_{mm,b} - r_{em,b} r_{me,b}) - 1. \end{aligned} \tag{4.35}$$

For constant ω, the propagation constants k_z which satisfy Equation 4.35 are the propagation constants of the eigenmodes.

Partial derivatives of Equation 4.35 may be found and inserted into Equations 4.5-4.7 in order to calculate the derivatives of k_z. In cylindrical coordinates with $m \neq 0$, finding derivatives of the Fresnel coefficients (r_{ee}, r_{em}, r_{me}, and r_{mm}) is not straightforward because a simple, closed-form expressions for the coefficients do not exist. Instead, they are generated by solving Equation 2.83, repeated here symbolically:

$$\Gamma^{(0)} = (M^{(0)})^{-1} V^{(0)}. \tag{4.36}$$

Derivatives of Equation 4.36 are taken and listed below. The equations generated are solved to generate the corresponding derivatives of r_{ee}, r_{em}, r_{me},

and r_{mm}.

$$M\Gamma^{(1)}_{\omega} = M^{(1)}_{\omega}\Gamma^{(0)} \tag{4.37}$$

$$M\Gamma^{(2)}_{\omega\omega} = M^{(2)}_{\omega\omega}\Gamma^{(0)} + 2M^{(1)}_{\omega}\Gamma^{(1)}_{\omega} \tag{4.38}$$

$$M\Gamma^{(3)}_{\omega\omega\omega} = M^{(3)}_{\omega\omega\omega}\Gamma^{(0)} + 3M^{(2)}_{\omega\omega}\Gamma^{(1)}_{\omega} + 3M^{(1)}_{\omega}\Gamma^{(2)}_{\omega\omega} \tag{4.39}$$

$$M\Gamma^{(1)}_{k_z} = M^{(1)}_{k_z}\Gamma^{(0)} \tag{4.40}$$

$$M\Gamma^{(2)}_{k_zk_z} = M^{(2)}_{k_zk_z}\Gamma^{(0)} + 2M^{(1)}_{k_z}\Gamma^{(1)}_{k_z} \tag{4.41}$$

$$M\Gamma^{(3)}_{k_zk_zk_z} = M^{(3)}_{k_zk_zk_z}\Gamma^{(0)} + 3M^{(2)}_{k_zk_z}\Gamma^{(1)}_{k_z} + 3M^{(1)}_{k_z}\Gamma^{(2)}_{k_zk_z} \tag{4.42}$$

$$M\Gamma^{(2)}_{k_z\omega} = M^{(2)}_{k_z\omega}\Gamma^{(0)} + M^{(1)}_{k_z}\Gamma^{(1)}_{\omega} + M^{(1)}_{\omega}\Gamma^{(1)}_{k_z} \tag{4.43}$$

$$M\Gamma^{(3)}_{k_zk_z\omega} = M^{(3)}_{k_zk_z\omega}\Gamma^{(0)} + M^{(2)}_{k_zk_z}\Gamma^{(1)}_{\omega} + 2M^{(2)}_{k_z\omega}\Gamma^{(1)}_{k_z} + 2M^{(1)}_{k_z}\Gamma^{(2)}_{k_z\omega} + M^{(1)}_{\omega}\Gamma^{(2)}_{k_zk_z} \tag{4.44}$$

$$M\Gamma^{(3)}_{k_z\omega\omega} = M^{(3)}_{k_z\omega\omega}\Gamma^{(0)} + M^{(2)}_{\omega\omega}\Gamma^{(1)}_{k_z} + 2M^{(2)}_{k_z\omega}\Gamma^{(1)}_{\omega} + 2M^{(1)}_{\omega}\Gamma^{(2)}_{k_z\omega} + M^{(1)}_{k_z}\Gamma^{(2)}_{\omega\omega}. \tag{4.45}$$

4.2 Lossy Waveguides, Plasmons, and Resonators

In lossy resonators, individual modes maintain their shape as they propagate, but they decay exponentially in amplitude as they travel through the structure. There are two possible mechanisms which will result in a mode being lossy: First, one or more layers may be made of an intrinsically lossy material, as discussed in Section 2.2.2. Second, if no layer has an index of refraction higher than that of the surrounding medium, then light will not be trapped in the waveguide by total internal reflection. These modes, known as leaky modes, are actually another way of looking at free-propagating modes which scatter from the structure.

The mathematical manifestation of a lossy mode is a complex propagation constant $k_z = k'_z + jk''_z$. A complex refractive index will result in a field which decays exponentially with a length constant $L = 2k''_z$. The exponential decay is expected as a generalization of Beer's law.

Even with lossy modes, the round trip phase shift must still be 2π, so Equation 4.4 remains valid. However, locating the roots of f by sampling will not be successful. First, the search region is two-dimensional, comprising the real and imaginary parts of the propagation constant k_z (k'_z and k''_z). The region will be computationally expensive to sample. Secondly, one

cannot determine where f changes sign. The Delves-Lyness algorithm [32] uses Cauchy's argument principle and numerically calculated path integrals to locate zeros in a closed region in the complex plane. This algorithm is described in Appendix D. The use of this algorithm for finding roots was first suggested by Chilwell and Hodgkinson [8] in 1984, and was used with the traditional transition matrix algorithm.

4.2.1 Plasmonic Waves

A special type of guided wave, known as a plasmon, occurs at interfaces between dielectrics and metals. The wave propagates along the interface and is evanescent (exponentially decaying) in both materials. Physically, a plasmon is a state which exists as a superposition between a photon and a charge oscillation. Mathematically, the wave exists because the relative permittivity of metals is negative.

The dispersion relationship for a plasmon is known to be:

$$k_z = \frac{\omega}{c}\sqrt{\frac{\epsilon_d \epsilon_m}{\epsilon_m + \epsilon_d}}, \tag{4.46}$$

where ϵ_d and ϵ_m are the permittivities of the dielectric and metals, respectively. Plasmon modes may only exist when the real part of ϵ_m is negative and $-\Re\{\epsilon_m\} > \epsilon_d$. For modes defined by a single air-metal interface, only TM modes exist. Both TE and TM modes may exist in structures comprising many layers.

Formulas have been developed for plasmonic structures with up to three layers [5] but, in general, no simple analytic solution exists for plasmon modes with an arbitrary number of layers. The strategy for locating plasmon modes will be to transform the plasmonic problem into an ordinary guided wave problem. To do so, an imaginary zero-thickness, high-index layer is placed inside the structure. Because the layer has zero thickness, it will not affect the mode. In practice, all metals are lossy at optical frequencies and so it is necessary to use the Delves-Lyness algorithm to locate the propagation constants. Exact values for dispersion may be determined with the same methods as those for ordinary guided modes (Equations 4.5-4.7). Although we have shown this to be true for the planar case, the same mathematics will

work to find plasmonic modes in layered cylindrical and spherical structures.

4.2.2 Resonator Modes

A resonator is a structure which confines light. In practice, the confinement is temporary, as light is lost either from absorption inside the structure or transmission out of the structure. Like leaky modes, resonator modes are not a fundamentally different type of mode, but rather are simply another way of looking at the scattering of light by a structure. The resonator problem is to find a frequency ω such that the magnitude and phase of a wave are unchanged by one round-trip through the resonator. Because the confinement is temporary, the frequency must be complex so the field decays inside the resonator as a function of time.

Resonator modes in planar structures occur at interfaces normal to the boundaries between layers. If the fields did not travel perpendicular to the surface, the power would 'walk off' any finite-sized structure. The field must be single-valued and so the eigenequation is:

$$r_a(\omega)r_b(\omega) = 1. \tag{4.47}$$

Equation 4.47 is similar to Equation 4.3, but instead of solving for k_z, we are solving for ω. In using the Delves-Lyness algorithm, we integrate in complex ω space.

A simple resonator is formed by a planar slab of dielectric, surrounded by two other materials. The reflection coefficients in this case are simply the Fresnel reflection factors, which will be called r_1 and r_2. The roots of Equation 4.47 are:

$$\omega'_m = \frac{\pi c}{nL} m \tag{4.48}$$

$$\omega'' = \frac{c}{2nL} \ln |r_1 r_2|, \tag{4.49}$$

where r_1 and r_2 are the two reflection coefficients, and L is the length of the cavity.

A figure of merit for resonators is the quality factor, or Q-factor, of the mode. The Q-factor is defined as $Q = \omega'/\omega''$. It has several interpretations, including: 1) being the ratio of the width of the resonance to the center

frequency and 2) the number of oscillations which take place before the amplitude is damped to $1/e$ of their original amplitude.

4.2.3 Whispering Gallery Modes

Spheres possess a special class of resonator modes known as whispering gallery modes (WGMs). Traditionally, the modes are viewed as light being guided around the circumference of the sphere by total internal reflection. There are two types of whispering gallery modes, TE^r and TM^r. The procedure for finding the modes is the same.

To find the whispering gallery modes with the phase method, we can view the sphere like any other resonator, with waves propagating in the $+\hat{r}$ and $-\hat{r}$ directions. For a simple sphere, a high-index, zero-thickness layer is inserted immediately outside the sphere's surface. Equation 4.47 is still the eigenequation, but with r_a interpreted as the inward reflection coefficient and r_b interpreted as the outward reflection coefficient from within the shell. Because the confinement of light is only temporary, the frequencies which satisfy the eigenequations must be complex, and so the Delves-Lyness algorithm must be used to solve the problem. However, the theoretical loss from spheres is extremely high ($>10^{20}$) and in practice is limited by scattering from within the material.

The traditional method for finding the eigenequation for whispering gallery modes of a sphere is as follows: Because ℓ and m are conserved in structures of circular symmetry, only the r-component of the field should be considered. Because the inside of the sphere contains the origin, the field must take the form $\hat{J}_\ell$. Outside the sphere, there may exist only a field propagating away from the sphere, so the field must be of the form$\hat{H}_\ell^{(1)}$. If the sphere has radius a, then by enforcing boundary conditions at the sphere-air interface, the eigenequations are found to be:

$$\frac{\hat{J}'_\ell(k_1 a)}{\hat{J}_\ell(k_1 a)} = \frac{n_2}{n_1}\frac{\hat{H}_\ell^{(1)\prime}(k_2 a)}{\hat{H}_\ell^{(1)}(k_2 a)} \tag{4.50}$$

for TE modes and

$$\frac{\hat{J}'_\ell(k_1 a)}{\hat{J}_\ell(k_1 a)} = \frac{n_1}{n_2}\frac{\hat{H}_\ell^{(1)\prime}(k_2 a)}{\hat{H}_\ell^{(1)}(k_2 a)} \tag{4.51}$$

for TM modes.

4.3 Implementation and Validation

Programs implementing the various techniques discussed in this chapter were coded in MATLAB. Both the algorithms and the code implementing them are complex. To show that neither the theory nor the code are in error, the results of the Airy-Markov calculation are compared with analytic solutions. In general, such solutions exist only in simple form for very simple systems. While the methods of this chapter will be validated against these simple, analytic solutions, they are capable of solving problems with many more boundaries. By creatively constructing test cases, the tests can be much more general than they first appear. For brevity, only a sample of the test cases are discussed.

Figure 4.2 shows the result of a validation test. The TE modes of a 250 nm thick planar waveguide, made of sapphire and surrounded by air on both sides, were calculated by finding the roots of Equation 4.3. The solid lines represent the propagation constants and the derivatives calculated with an analytic formula. The red line represents the fundamental (zero nodes) mode while the blue line is for the first excited (one node) mode. The blue mode disappears around 720 nm because the first mode cutoff lies at longer wavelengths (it becomes a free-propagating mode). The black dots are the values calculated with the Markov-Airy method and the agreement is excellent. The maximum value for the relative error between the two methods is 10^{-14} for k_z, 10^{-12} for k_z', 10^{-10} for k_z'', and 10^{-8} for k_z''', where primes represent total derivatives with respect to ω.

The modes of a step cylindrical waveguide were also computed and compared with an exact, analytical solution [30]. With zero-thickness shells placed inside and outside of the central core, the Markov-Airy method agreed with the analytic solution with a relative error of around 10^{-14}.

Figure 4.3 shows an example of locating the modes of a resonator with the Delves-Lyness algorithm. The resonator comprised a 1 μm thick slab with an index of refraction of n = 2.4, and surrounded by air. The Delves-Lyness algorithm was used to search for modes between 400 nm and 1200 nm. Panel (a) shows the analytic locations of the roots with green x's and the positions

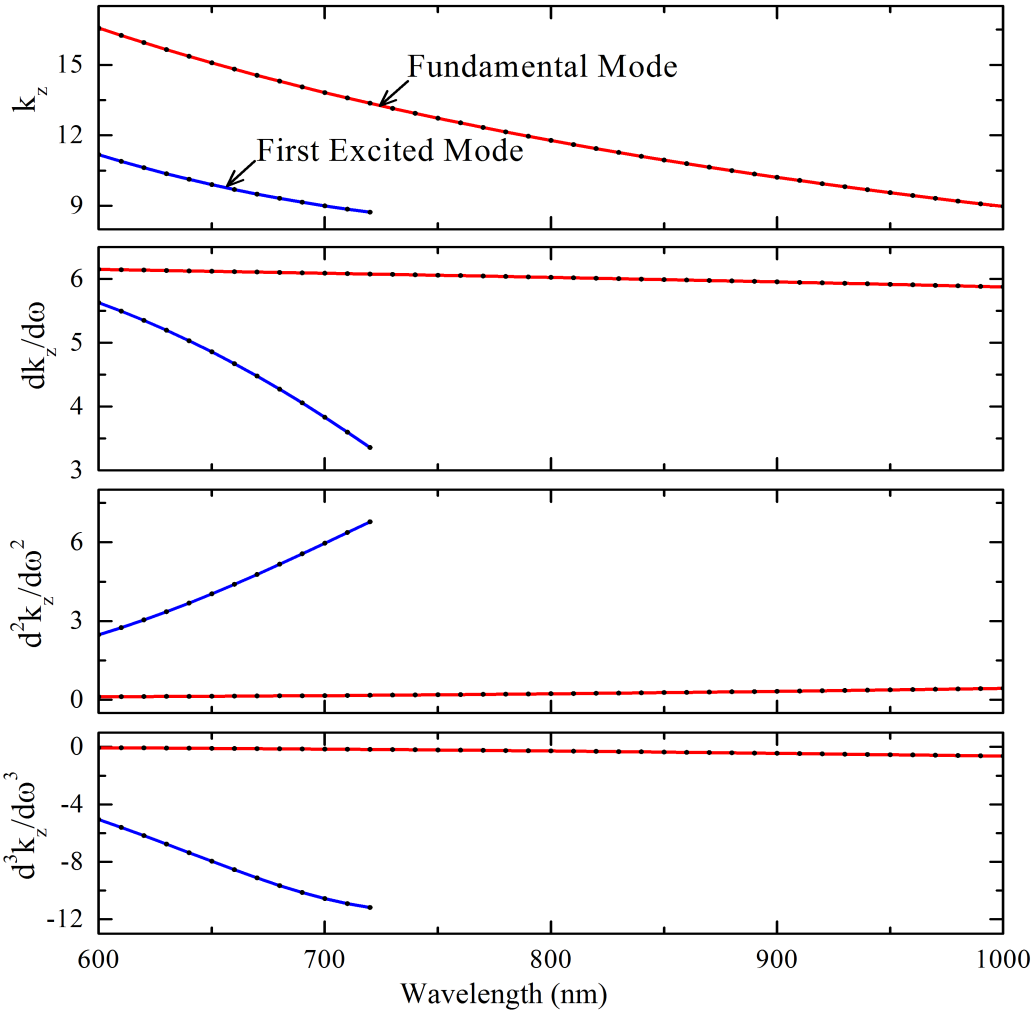

Figure 4.2: Plots of the dispersion of a 250 nm thick sapphire waveguide. The units of the four plots from top to bottom are rad/μm, rad-fs/μm, rad-fs^2/μm, and rad-fs^3/μm, respectively.

calculated with the Markov-Airy method with blue circles. Note the scale on the y axis. In an effort to illustrate the accuracy of this computation in calculating the root locations, the relative error in the magnitude is shown in panel (b).

In testing various parameters of the cavity for Figure 4.3, it was found that the implementation of the algorithm was unable to correctly calculate Q-factors over 10^{15}. The maximum Q-factor which could be calculated was limited because the algorithm was implemented with standard double floating point arithmetic, which is accurate to only 16 decimal digits. Greater precision could be obtained by implementing the algorithm with arbitrary-

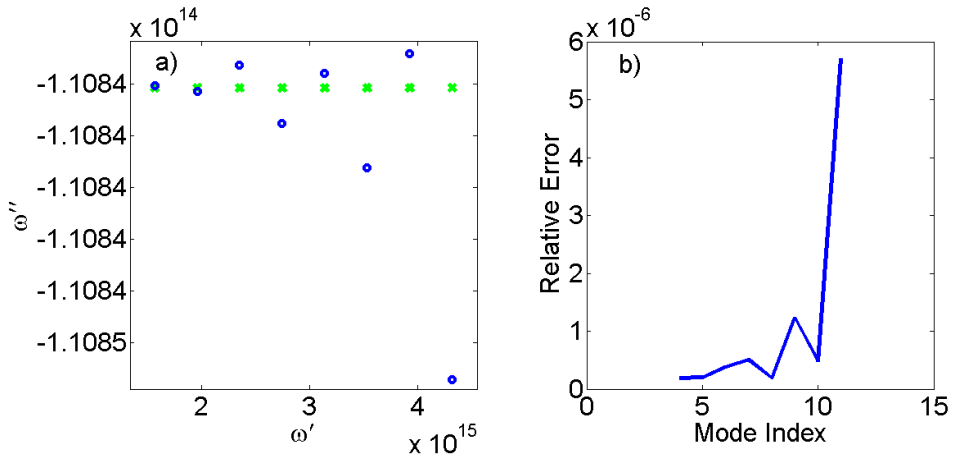

Figure 4.3: (a) Location of resonances of a 1 μm thick slab resonator with a refractive index of 2.4. The analytic positions are marked with green x's while the calculated positions are marked with blue circles. (b) Relative error in the magnitude of the calculated position of the resonances, as compared to the exact value.

precision arithmetic.

4.4 Conclusion

The problem of locating waveguide and resonator modes of layered structures has been discussed. The eigenequations for all of the problems discussed here take on a similar form (Equation 4.3 or 4.47). One notable exception is finding the waveguide modes of a cylindrical resonator for which $m \neq 0$. In that case, the eigenequation is that of Equation 4.35.

Arbitrary waveguide and plasmon modes have been calculated previously with the T-matrix method [8, 9]. Because the method discussed in this chapter is based on the Airy-Markov method, we would expect a benefit of improved numerical stability. Additionally, this chapter is the first to present exact formulas for the derivatives of k_z with respect to ω. These derivatives are expected to be of value when studying the interaction of plasmons with ultrafast laser pulses.

CHAPTER 5

MICROSPHERE-STABILIZED PLANAR RESONATORS

5.1 Introduction

Researchers have long studied the whispering gallery modes of millimeter and micron-sized spheres [33]. In this work an alternative method to creating a resonator using microspheres is explored. Specifically, spheres are placed inside a cavity formed by two planar mirrors, as seen Figure 5.1. A quantum dot solution surrounds the sphere and acts as a gain medium. The spheres can act inside the cavity as thick lenses. If the mirrors are sufficiently close to one another, the spheres will stabilize the cavity and lower its threshold. Conversely, when the mirror separation increases beyond a critical value, an unstable cavity is formed which increases the threshold pump energy.

There are numerous potential applications for this system. By scanning the mirror separation while pumping the quantum dots, it may be possible to determine the distribution of sphere sizes by noting at what distances the spheres stop lasing. Because of the low lasing threshold, large areas may be pumped and imaged at once, which will enable the sphere size statistics to be rapidly determined. Cytometry may be an additional application. A fixed-mirror system has been shown to be capable of resolving morphology in red blood cells [34]. Applications in sensing may also be envisioned.

5.1.1 Experimental Details

The optical cavity was formed by two high-performance dielectric mirrors manufactured by CVI. They were rated for $> 99.97\%$ reflectivity at 633 nm at normal incidence by the manufacturer. The actual reflectivity around 660 nm (where lasing was observed) was smaller both because the peak wavelength of the gain spectrum was away from the design wavelength of the mirror

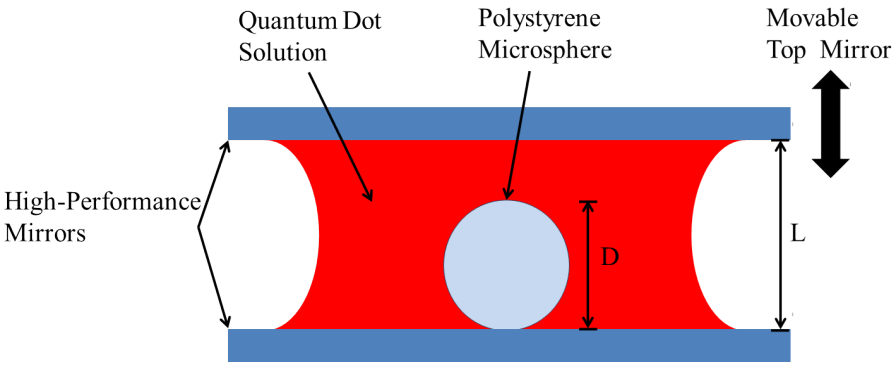

Figure 5.1: Diagram of cavity, not to scale. The transverse size of the droplet was much larger than an individual sphere.

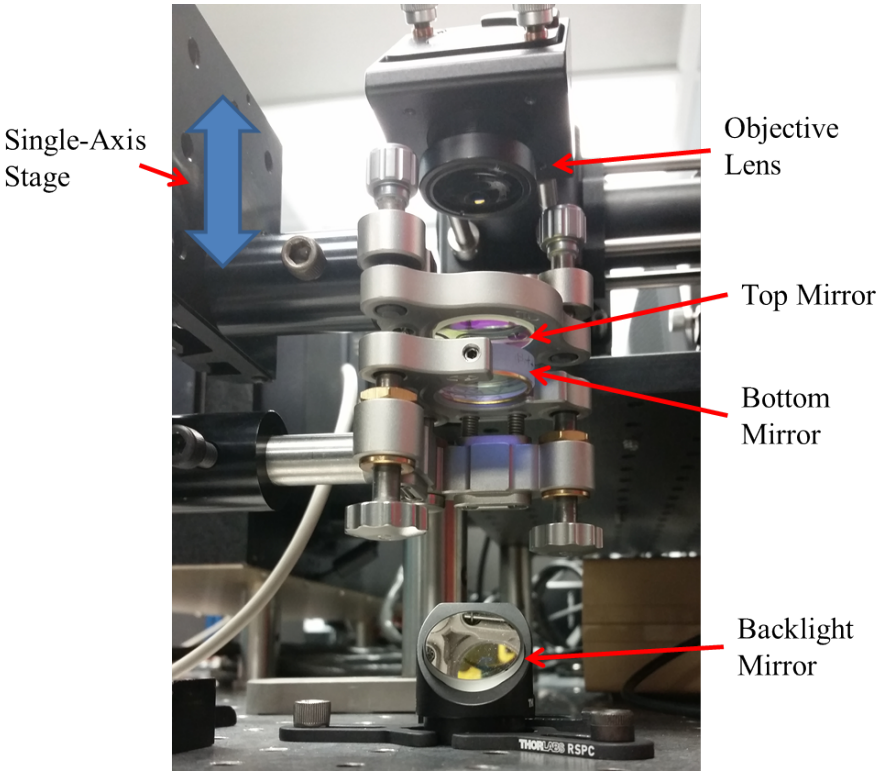

Figure 5.2: Photograph of cavity. The mirrors were brought apart so they could be seen and needed to be much closer together for experiments. The LED used for illumination has been removed for this picture.

and because exposure to water likely lessened the reflectivity of the mirror. Reflectivity at the pump wavelength of 532 nm was approximately 40%. Lower reflection at the pump wavelength would have been preferable, but these mirrors could be purchased "off-the-shelf".

Polystyrene microspheres were purchased from Spherotech. Three different

diameter spheres were used: 75, 100, and 196 μm. The sphere diameters varied approximately 5% around the mean value. They were shipped suspended in water and with their surface functionalized with streptavidin. However, nothing was bound to them during the course of experiments.

Quantum dots were chosen as the gain medium for the system. An 8 μM solution (Qdot 655 ITK) of quantum dots was purchased from Life Technologies. The fluorescence spectrum was centered at 655 nm with a full width at half maximum of 30 nm. A coating of carboxyl allowed the quantum dots to be water soluble. There were several reasons for choosing quantum dots instead of dyes. The primary reason is that dye degrades when it is strongly pumped. After pumping, some dye molecules will be trapped in the metastable triplet state, while other dye molecules will be rendered permanently optically inactive [35]. In dye lasers, the problem is solved by circulating the dye, but in this system with $\sim$100μm mirror separation, circulation was impractical. Quantum dots do not suffer from photodegradation. Secondly, dyes, with their relatively narrow absorption band, would have required customized mirrors for this experiment. In contrast, quantum dots may be pumped at practically any wavelength shorter than their emission band, which greatly relaxed the requirements for the mirrors.

Figure 5.2 shows a picture of how the cavity was mounted. The two mirrors were mounted in stainless-steel optical mounts. The top mirror was mounted on Newport 421 series linear translation stage with a micrometer drive. This arrangement allowed the length of the cavity to be changed without (to first order) affecting the focus. The stage was positioned so that both its internal spring and gravity pulled downward together against the micrometer drive. As gravity did not work to cancel the force of the spring, this made the mounting more stable. The linear stage and the bottom mirror were in turn mounted on a three-axis stage. The pump laser and the view from the microscope were fixed, so the three-axis stage was used to bring the mirrors into focus. The spheres were illuminated from below by a collimated LED.

A microscope was constructed to pump the spheres and analyze the light emitted by them. The design of the microscope is shown in Figure 5.3. The objective was a simple, plano-convex lens made with BK7 glass. It had a focal length of 50 mm. The numerical aperture was 0.25, which gives a visible light resolution of approximately 3.2 μm. After passing through the dichroic, light was focused by a tube lens with a 200 mm focal length. After the tube lens,

the light passed through a band-stop filter to remove scattered pump light and was incident on a 50/50 beamsplitter, which sent half the light onto a CCD camera and the other half onto the entrance slit of a spectrometer. The overall magnification of the system is 4X. The image of the smallest resolvable feature was then 12.8 μm. Because the pixel size on the camera was 4.65 μm, the resolution was limited by the optical system and not the CCD pixel size. Optics in the lower two beam paths of the diagram are used to align, control, and measure the pump light. Kinematic mirror #1 allows easy switching between the helium-neon (HeNe) laser for alignment and the Nd:YAG laser for experiments. A neutral density filter-wheel allows the power to be controlled. It has six filters of optical density 0, 0.1, 0.2, 0.4, 0.6, and 1.0. When the kinematic mirror #2 is removed, the pulse energy may be measured. The center beam path is used for alignment and the addition of lenses to focus or defocus the pump light.

The pump source for these experiments was a Spectra Physics Quanta Ray Pro at 532 nm, attenuated to microjoule pulse energies. Pulses from this laser had a pulse width of approximately 10 ns FWHM. Between the mirrors of the cavity, the spot size of the pump had a diameter of approximately 75 μm. The spectrometer was a Princeton Instruments Acton 2750 0.75 m imaging spectrometer with a PI-MAX4 detector. The PI-MAX4 is an intensified CCD (ICCD) detector, which is a CCD with a microchannel plate amplifier mounted on top. Using voltage for control, gains of up to 100x can be achieved with time resolution as small as 2.5 ns. The gain increased the sensitivity, while gating reduced the noise by only recording when signal is present.

5.1.2 Experimental Procedure

For alignment, the 550 nm longpass dichroic mirror was replaced by a 650 nm longpass dichroic to allow for the use of a HeNe laser for alignment. The bottom mirror was installed first and adjusted to retroreflect the HeNe laser. The top mirror was then installed and also adjusted to retroreflect the alignment laser. Based on the distance from the mirrors to the alignment aperture, the alignment error is estimated to be 3.4×10^{-4} rad.

Separately, 10 μL of a quantum dot solution and 1 μL of a sphere-containing

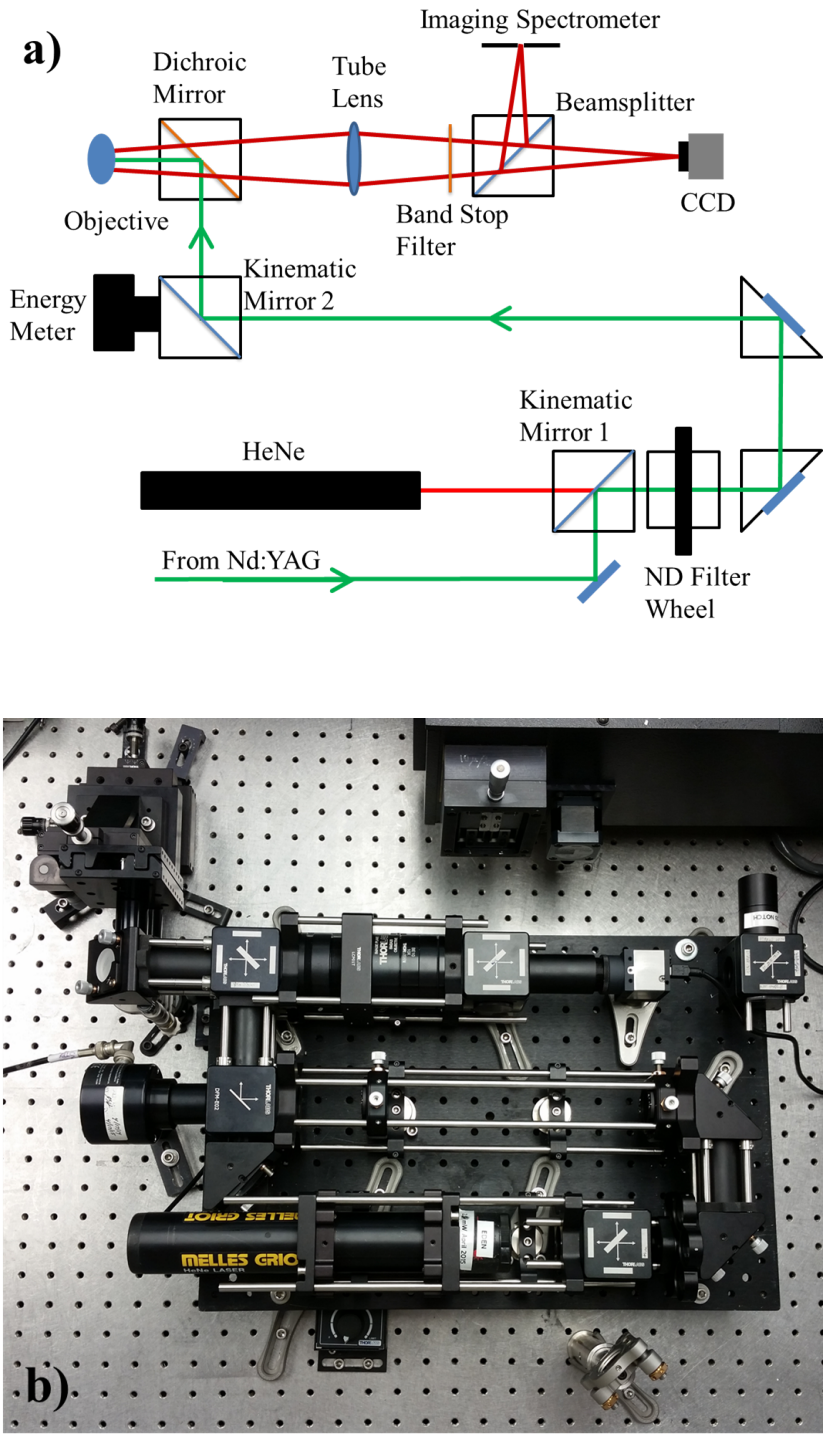

Figure 5.3: (a) Diagram of microscope. The objective lens points into the plane of the diagram. (b) Overhead view of actual microscope.

fluid were pipetted onto the bottom mirror. The spheres were shaken before pipetting to place them into suspension. The micrometer controlling the

stage on which the top mirror was mounted was then adjusted until the top mirror came into contact with the droplet. A change in the transverse size of the droplet was clearly visible, as the water adhered to both of the mirrors. The imaging lens was installed and the YAG laser turned on. The alignment of the mirrors could be further optimized by adjusting the top mirror to maximize the intensity of the lasing. In practice, however, only very small adjustments were necessary.

The relationship between the micrometer reading and the actual cavity length was not repeatable after a new sample was loaded. This was because the adjustment of the mounts to make the mirrors parallel to each other altered the distance between the surface of the mirror and the post to which the mount is attached. Instead, the distance between the mirrors was inferred from the wavelength spacing between adjacent longitudinal modes (known as the free spectral range) in the cavity. The relationship between the free-spectral range and the mirror spacing is given by:

$$FSR\ (\text{nm}) = \frac{\lambda_0^2}{2nL}, \tag{5.1}$$

where λ_0 is the center wavelength of the lasing, n is the refractive index of the material between the cavity, and L is the length of the cavity. Equation 5.1 is only valid when no sphere is present.

After the mirrors were aligned, the last step before beginning experiments was to establish a minimum mirror separation. The pump was focused onto an area of the quantum dot solution with no spheres. The spheres were then carefully brought together while monitoring the free spectral range with the spectrometer. For each sphere size, a maximum value for the free spectral range, above which the mirrors would be closer together than the sphere diameters, was calculated. The top mirror was lowered until the free spectral range was roughly 10% smaller than the maximum value. The reading on the micrometer was noted and the experiments could begin. In data runs where the cavity length was being scanned, the top mirror was always scanned up. This was to eliminate any effects from backlash in the micrometer screw.

After each set of experiments, the spheres and quantum dots were rinsed off the mirrors into a disposal container. The mirrors were then drag-and-drop cleaned with lens paper, at least twice with water and twice with ethanol. The mirrors were then covered with lens paper and left to dry on a gently

heated hotplate (30 °C).

5.2 Data and Analysis

5.2.1 General Characteristics

Depending on the mirror separation, spheres have been observed to both lower and raise the lasing threshold. As will be seen below, when the mirrors are sufficiently far apart, the cavity formed with the sphere acting as a lens becomes unstable and the modes suffer from high diffraction loss. The modes of the planar cavity, however, are always marginally stable for any mirror separation. In these situations ,the lasing is observed in between spheres, but when the pump is brought over a sphere, the lasing shuts down. However, when the mirrors are close together to form a stable cavity, the threshold is dramatically lowered. Particularly with the 75 μm spheres with mirror separations under 100 μm, it can be difficult to make a sphere-free area to lase, while the spheres themselves lase easily.

Threshold pump energy was observed to be less than 0.1 μJ. Even at these pump energies, the pulse still has sufficient energy to excite every quantum dot in the pump volume. Pump pulse energies above 3 μJ generally resulted in the spheres being damaged. Optical damage of the polystyrene spheres from their own lasing modes was likely the cause. If the spheres were prevented from lasing by having the mirrors far apart, then they could withstand intense pump light, then lase again when the spheres were moved back together. The approximate diameter of the lasing mode inside the sphere was 10 μm.

Under excitation, the spheres generally appear as bright, homogeneous circles, as shown in the upper left image in Figure 5.4. The uniformity is the result of the spheres lasing in a superposition of transverse modes. Sometimes, the spheres could be seen lasing in a single transverse mode. The remaining pictures which make up Figure 5.4 show the same sphere at different times. Possibly due to pump fluctuations, this particular sphere would rapidly switch among transverse modes. These images demonstrate that Gaussian modes are being observed from the spheres.

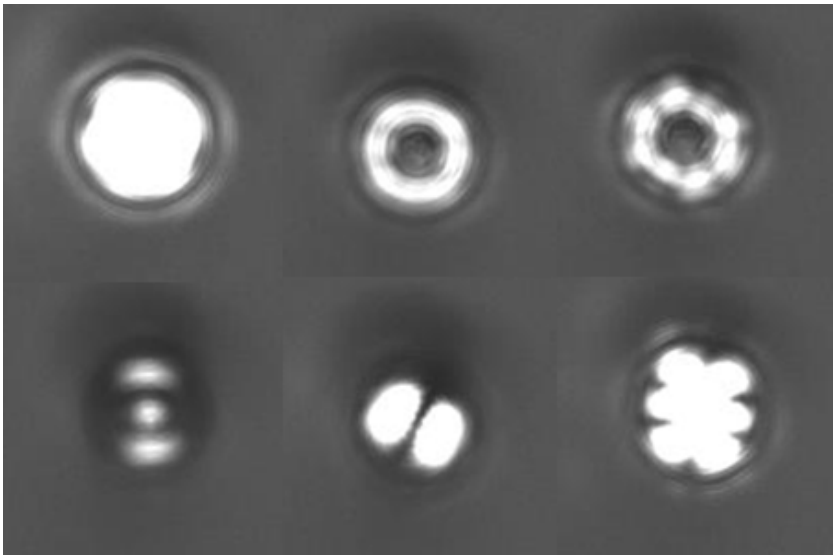

Figure 5.4: Gaussian modes observed in a 75 μm sphere. Clockwise from top left: An incoherent sum of modes, LG_{10} ('donut' mode), LG_{04}, TEM_{12}, TEM_{10}, TEM_{02}.

5.2.2 Spectral Changes

As compared to a planar cavity, the presence of a sphere fundamentally modifies the lasing spectrum in two ways: First, it is well known that the transverse modes of planar resonators are frequency degenerate [36, 37]. However, higher-order transverse modes of non-planar resonators are frequency shifted with respect to the TEM_{00} mode. Therefore, the spectrum of the cavity with a sphere is much richer in lines. Secondly, the sphere ($n_s = 1.58$) displaces water ($n_w = 1.33$) in the cavity. The free spectral range (the distance between adjacent longitudinal modes) would be expected to decrease.

The free spectral range with a sphere present may not be calculated by simply replacing nL with the new optical thickness in Equation 5.1. The change in the dielectric media at the sphere/water boundary creates internal resonances in the cavity, which must be taken into account. The theory described in Chapter 4 may be used to model the cavity. Assume the reflection coefficient of the mirrors to be 1, for simplicity. Then, consider a point just inside the sphere at the sphere/water boundary. The reflection coefficients looking back into the sphere and into the water are

$$r_s = e^{j2k_sD} \tag{5.2}$$

and

$$r_w = \frac{r_{sw} - e^{j2k_w(L-D)}}{1 + r_{sw}e^{j2k_w(L-D}}, \tag{5.3}$$

respectively. Here r_{sw} is the reflection coefficient of just the sphere/water interface (starting in the sphere), k_s is the magnitude of the propagation vector in the sphere, k_w is the magnitude of the propagation constant in water, D is the diameter of the sphere, and L is the total separation between the mirrors. In order for the field to be single valued, the product of these reflection coefficients must be 1. Therefore, the eigenequation for the resonances is:

$$e^{j2k_sD}\frac{r_{sw} - e^{j2k_w(L-D)}}{1 - r_{sw}e^{j2k_w(L-D}} = 1. \tag{5.4}$$

Equation 5.4 can be solved numerically to determine the free spectral range. Figure 5.5 shows a change in the free spectral range from 520 GHz to 440 GHz when a 200 μm sphere is placed between two mirrors 216 μm apart, which is in agreement with theory.

In addition to the shrinking of the free spectral range, Figure 5.5 also shows additional transverse modes, as predicted at the opening of this section. A wider view of the change in the spectra with a different sphere is shown in Figure 5.6. Once again, extra lines due to the transverse modes are present. Analysis of the lines between 656 nm and 660 nm shows the expected shift in the free spectral range from 800 GHz to 720 GHz.

5.2.3 Stability

A key property of laser resonators is stability. From a geometrical optics point of view, a ray of light inside a stable cavity will stay within a finite distance from the optical axis as it is reflected between the end mirrors of the cavity. In an unstable cavity, any ray would go arbitrarily far from the origin as it bounces between the end mirrors. A planar cavity, in which the cavity comprises two planar mirrors with a uniform material in between, is known to be 'critically stable.' While rays perfectly normal to the mirror surface will not walk off, any other rays will walk away from the origin.

In the wave picture of light, only stable cavities support finite-sized Gaussian modes [38]. These modes have low diffraction loss because the light is confined around the optical axis and therefore traverses the pumped region

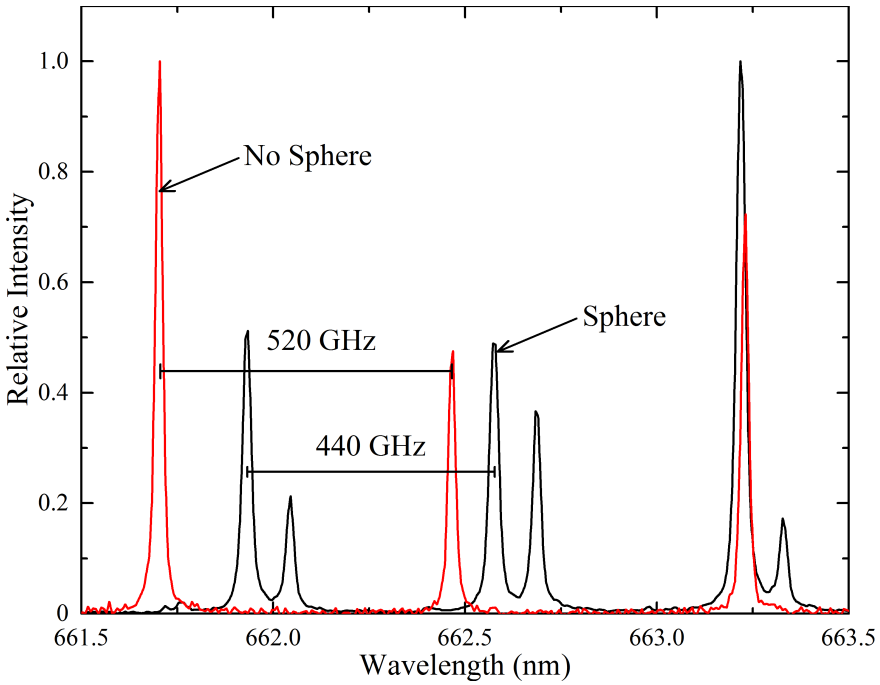

Figure 5.5: Change in spectrum observed with a 200 μm diameter sphere and the mirrors separated by approximately 216 μm.

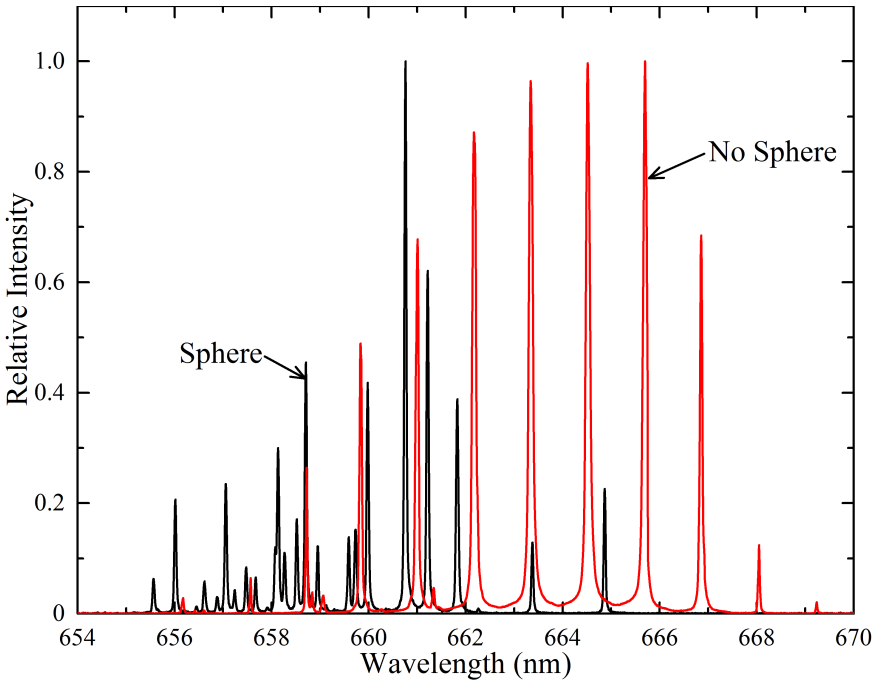

Figure 5.6: Spectra observed from and a 140 μm planar cavity when the pump beam excites just the cavity and when it strikes a 75 μm diameter sphere in the cavity.

Table 5.1: ABCD matrices of the elements inside the cavity.

P_r	$\begin{pmatrix} 1 & r \\ 0 & 1 \end{pmatrix}$	Propagate distance r
P_w	$\begin{pmatrix} 1 & w \\ 0 & 1 \end{pmatrix}$	Propagate distance w
P_{cc}	$\begin{pmatrix} 1 & 0 \\ \frac{n_w - n_s}{R n_w} & \frac{n_s}{n_w} \end{pmatrix}$	From sphere into water
P_{cv}	$\begin{pmatrix} 1 & 0 \\ \frac{n_w - n_s}{R n_s} & \frac{n_w}{n_s} \end{pmatrix}$	From water into sphere

many times. In contrast, light in an unstable resonator will pass through the pumped region only a few times before exiting the resonator. Diffraction losses increase as the parameters of a stable cavity are altered to bring it closer to the limit of stability [36, 39]. Planar resonators, being on the border between stable and unstable, suffer from high diffraction loss.

Without the microspheres, the cavities under observation would be planar cavities. The spheres themselves act as thick lenses which, under the appropriate circumstances, focus diffracting light and stabilize the cavity. Cavity stability can be analyzed with ABCD matrix theory [38, 37]. ABCD matrix theory relates the distance and slope with respect to the optical axis at one point in a paraxial optical system to another point. A list of the ABCD matrices of the individual elements necessary to analyze the cavity is given in Table 5.1. To determine stability, it is necessary to find the net ABCD matrix for one round-trip through the cavity. The round-trip ABCD matrix, starting from the center of the sphere and proceeding upwards, is:

$$P_{rt} = P_r P_{cc} P_w P_w P_{cv} P_r P_r P_{cc} P_{cv} P_r. \tag{5.5}$$

The condition for stability is that

$$0 \leq \frac{A + D + 2}{4} \leq 1, \tag{5.6}$$

where A and D are the upper left element and lower right elements of P_{rt}, respectively. The term in the center of the inequality in Equation 5.6 is known as the stability parameter.

A plot of the stability parameter as a function of the mirror separation for each of the three sphere sizes used in these experiments is shown in Figure

5.7. There is clearly a critical value of the mirror separation beyond which the cavity becomes unstable. To demonstrate this phenomenon, individual spheres were selected and their spectra observed as the mirrors were moved apart. The ∘'s on the graph indicate specific mirror separations where lasing was observed, while x's indicate mirror separations where lasing was not observed. Cutoff was observed for the 100 μm spheres, but the exact mirror separation at which it occurred was not recorded. The experiment was repeated with many more spheres than shown in the graph, but quantitative data were not recorded. Even with the pump energy increased by a factor of more than 30 over the observed threshold value for smaller mirror separations, the spheres could not be made to lase when the cavity length exceeded the critical value.

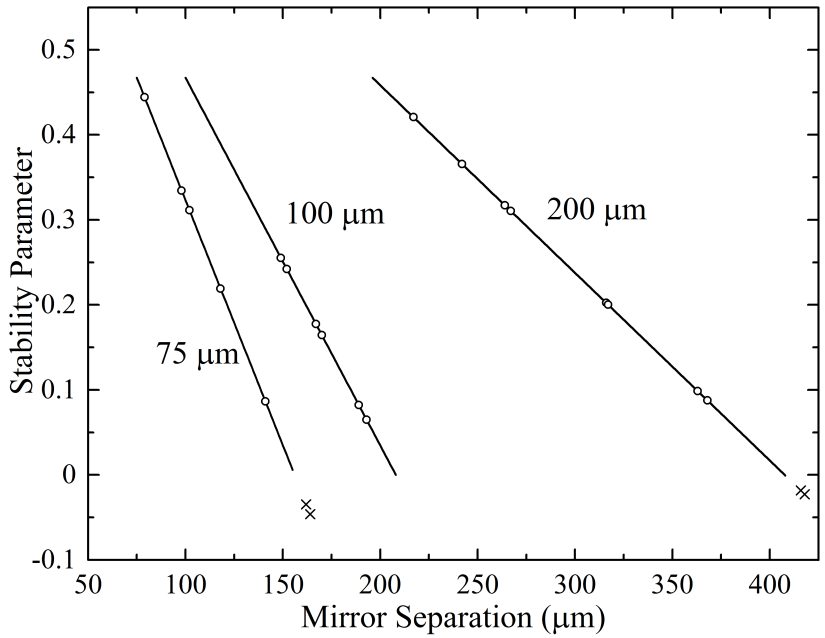

Figure 5.7: Plot of the stability factor (Equation 5.6) as a function of the mirror separation for the three different sphere diameters. The circles (∘) indicate datapoints where a sphere was observed to lase while x's indicate mirror separations where a sphere could not be made to lase.

5.3 Temporal Narrowing

The pulses of light coming from planar resonators with and without spheres were observed to be shorter than the pump pulse. If the light observed was

due to fluorescence, the output light would be at least as long as pump pulse, because fluorescence is a linear phenomenon. Therefore, the shortening of pump pulse is interpreted as a further indication of lasing. Two typical plots of intensity as a function of time are shown in Figure 5.8. In part (a), the light from the cavity with a sphere is observed to turn on later. The delay could represent the buildup time of the laser inside the cavity. Part (b) shows that the pulse from the planar cavity is shortened much more than if a sphere were present. Furthermore, intensity falls below that of the pump both before and after the peak. A possible explanation is that due to the much higher threshold of the planar cavity, it is only over threshold for a short time window around the peak.

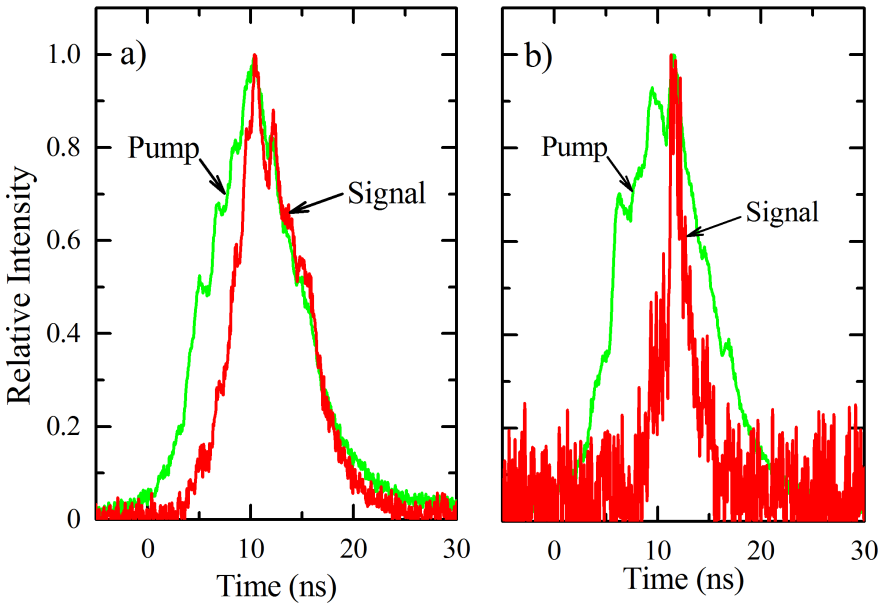

Figure 5.8: Plots of the intensity of light as a function of time when the pump is focused: (a) on a sphere or (b) no a sphere with 1.5 μJ pump pulses. All pulses have been normalized so their peak intensity is 1. The large noise in the signal plot on the right is because the intensity is much lower.

5.3.1 Laser Arrays

Sphere-stabilized planar cavities may be used to create large laser arrays. Self-assembly techniques [40, 41, 42] have long been used to create two-dimensional arrays of spheres. Large numbers of spheres can then be pumped

simultaneously, using only modest pulse energies (recall that the threshold for an individual sphere was observed to be < 0.1 μJ).

A simple array has been created for demonstration purposes. Many more spheres than usual (3 μL of solution vs the normal 1 μL) were added to a sample. The spheres were assembled by pulling the mirrors apart, which decreased the cross sectional area of the droplet. As the droplet shrunk, the spheres were held in the droplet by their adhesion to the water. The mirrors were then pushed back together, slowly, to prevent the flowing quantum dot fluid from moving the spheres. The pump laser was expanded on the sample with a diverging lens, which increased the pump area to roughly 1.2 mm in diameter. The pulse energy was increased to 19 μJ. The result is shown in Figure 5.9. The picture on the left shows the spheres with LED illumination. The picture on the right shows the same spheres with the pump laser on. The bright spots all indicate lasing spheres. Note the lack of lasing in areas in between spheres. Although the array of Figure 5.9 only shows approximately 80 spheres lasing simultaneously, there should be no difficulty in increasing the number of spheres dramatically.

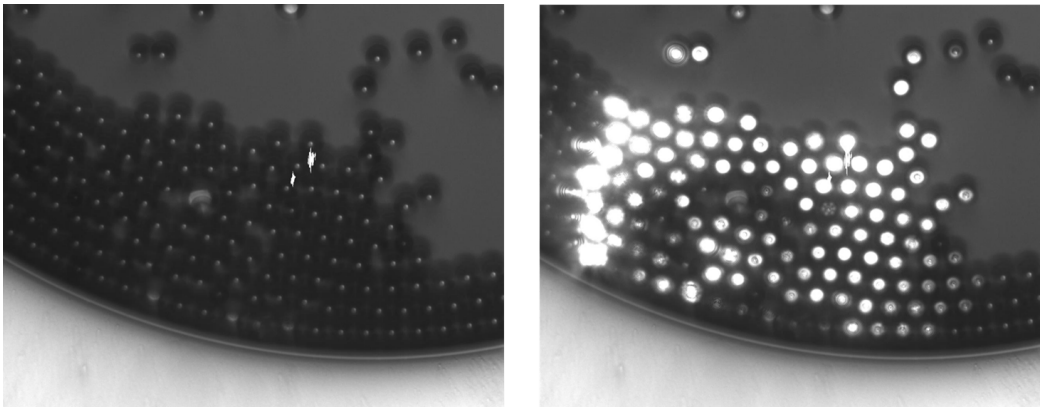

Figure 5.9: Left: An array of 75 μm spheres inside a planar cavity. Right: The same array of spheres lasing upon being pumped with a 19 μJ pulse. The bright spot above and to the right of the center is due to a damaged detector.

5.4 Conclusion

A new type of laser cavity, formed with planar mirrors stabilized by microspheres, has been demonstrated. The change in the spectrum of the planar

modes due to the presence of the sphere has been discussed and is in good agreement with theory. The maximum separation of the mirrors for the cavities to be stable has been calculated and the results agree with experiments. The lasing signal coming from the spheres has been observed to be temporally narrower than the pump pulse, but further investigation will be required to determine the exact cause. Finally, the ability to assemble the microspheres into an array of microlasers has been shown.

CHAPTER 6

CONCLUSION

A new method, based on Markov analysis, for calculating the interaction of light with layered materials has been presented. The method was developed in stages. First, the solutions for the fields in Cartesian, cylindrical, and spherical coordinates were derived. Then, the reflection and transmission coefficients at material boundaries were defined. Finally, the reflection and transmission coefficients were used in transition matrices in order to solve scattering, waveguiding, and resonator problems. Exact and numerically stable formulas for calculating dispersion were also developed.

Compared to the traditional method of solving these problems (the transmission matrix method) the Markov-Airy method is numerically more well-posed. For this reason, the new method is less likely to give an inaccurate result when calculations are performed with floating point arithmetic. The relative error in these calculations, as found by testing against analytic solutions, has been found to be on the order of machine epsilon (10^{-15}). The code implementing this functionality, including the validation files and examples, has been made freely available on a public repository.[1]

Specific problems solved in this dissertation include determining the reflectivity of a mirror, including the first, second, and third phase derivatives, which determine how a mirror will distort a short pulse. Furthermore, the problem of calculating the dispersion of planar waveguides, including plasmonic waveguides, has been solved in the general case. The exact formulas for calculating the derivatives of the propagation constant with respect to frequency are, to the author's knowledge, the first ever presented. Cylindrical waveguides were also discussed. Finally, planar resonators and whispering gallery mode resonators were discussed.

In the second section of this thesis, a new laser structure was demon-

[1] http://www.mathworks.com/matlabcentral/fileexchange/47360-markov-airy-mirror-zip

strated. Spheres placed between two planar mirrors acted as lenses which could stabilize or destabilize the cavity, depending on the mirror separation. The lasing threshold is raised or lowered correspondingly. The mirror separation at which the cavity became unstable was determined experimentally for several sphere sizes, and found to be in excellent agreement with theory. When the spheres made the cavity unstable, lasing did not occur at any pump energy.

Gaussian modes have been observed emanating from the sphere-stabilized cavities, both in the spatial profile of the beam and spectrally. Spatially, the spheres generally lase in a superposition of transverse modes, but at times they are observed lasing in a single transverse mode. By altering the lasing parameters, such as pump area or sphere size, it may be possible to force the spheres to lase in a single mode. Spectrally, the cavity always exhibits multiple longitudinal modes, as would be expected from the inhomogeneously broadened gain medium (quantum dots). When no sphere is present, the spectral lines are evenly spaced in wavelength. When spheres are present, many more lines are observed. The extra lines appear because transverse modes are not degenerate for Gaussian modes, whereas they are for a planar resonator.

The lasing signal has been observed to be temporally more narrow than the pump which pumps the cavity. This is interpreted as a sign of lasing. Modeling of the temporal narrowing and the output power was not possible because the mirrors have significant reflectivity at the pump wavelength. The pump light was, in effect, passing through a Fabry-Perot cavity. Small variations in the frequency spectrum of the pump laser could dramatically affect how much light was absorbed, an effect that could not easily be modeled. With new cavity mirrors which do not reflect at the pump wavelength, it should be possible to perform a more detailed analysis of the system.

The new structure could be used to study the spheres themselves, or by coating the spheres with binding sites, they could be used as sensors. The new method could also allow inexpensive manufacturing of large arrays of lasers with high mode quality.

APPENDIX A

MARKOV CHAINS IN PROBABILITY

A Markov chain is a model for a stochastic process in which populations are transferred among a finite number of states in discrete steps. The probability of transitioning to a specific state in the next step of the process depends only on the current state and does not depend on time.

The state of a Markov process can be represented by a column vector $\vec{v}$. The transition probabilities can be combined into a transition matrix $\mathbf{T}$. The element T_{mp} is the probability of transitioning from state p to state m in one step of the Markov process. If the initial distribution among states is given by v_0, the distribution after the first step of the Markov process is:

$$\vec{v}_1 = \mathbf{T}\vec{v}_0. \tag{A.1}$$

It can easily be seen that after k steps of the Markov process, the distribution among the states is:

$$\vec{v}_k = \mathbf{T}^k\vec{v}_0. \tag{A.2}$$

If a state returns to itself with unit probability and it is possible to arrive at that state from at least one other state, it is known as an absorbing state. The transition matrix of a Markov chain with L absorbing states and K non-absorbing (transient) states can be written (possibly with some reordering of the states) as:

$$\mathbf{T} = \left(\begin{array}{c|c} \mathbf{I} & \mathbf{C} \\ \hline \mathbf{0} & \mathbf{Q} \end{array}\right), \tag{A.3}$$

where $\mathbf{I}$ is a $L \times L$ identity matrix, $\mathbf{C}$ is $L \times K$ matrix, $\mathbf{0}$ is a $K \times L$ zero matrix, and $\mathbf{Q}$ is a $K \times K$ matrix.

Given an initial distribution of probability among the non-absorbing states, it is possible to calculate the final distribution of probability among the absorbing states. The idea is to find the expected number of visits to each

non-absorbing state. $\mathbf{Q}$ can be thought of as the transition matrix between the non-absorbing states. The probability of being in each state after k steps of the Markov process is given by the vector:

$$\vec{v}_k = \mathbf{Q}^k \vec{v}_0. \tag{A.4}$$

The expected number of visits to each state is the sum of the probability that a given state will be occupied at each step:

$$\begin{aligned} \vec{v} &= (\mathbf{I} + \mathbf{Q} + ... + \mathbf{Q}^n + ...)\vec{v}_0 && \text{(A.5)} \\ &= (\mathbf{I} - \mathbf{Q})^{-1} v_0 && \text{(A.6)} \\ &= \mathbf{N}\vec{v}_0, && \text{(A.7)} \end{aligned}$$

where $\mathbf{N} \equiv (\mathbf{I} - \mathbf{Q})^{-1}$.

The probability of settling in an absorbing state is the sum of the probabilities of being in each transient state at each step (the expected number of visits) multiplied by the probability of transitioning from that state to the absorbing state. The coupling from transient to absorbing states is encapsulated in the $\mathbf{C}$ matrix. Therefore, the final distribution of probability among the absorbing states is:

$$\vec{a} = \mathbf{C}\mathbf{N}\vec{v}_0, \tag{A.8}$$

where $\vec{a}$ is a $L \times 1$ column vector.

APPENDIX B

PHASE DERIVATIVES IN RADIAL COORDINATES

The reflection coefficient is a complex vector with both a magnitude and phase. The group delay, group delay dispersion, and third-order dispersion are critical concepts when considering broadband reflection. These quantities are the first, second and third derivatives, respectively, of the phase of the reflection coefficient with respect to angular frequency. In this appendix, it will be shown how to compute the phase derivatives of a complex function from the total derivatives.

B.1 One-Dimensional Phase Derivatives

As a complex number, the reflectivity may be expressed in terms of its magnitude and phase, which are both functions of ω as shown below:

$$\bar{R} = r(\omega)e^{j\theta(\omega)}. \tag{B.1}$$

The first, second, and third derivatives of Equation B.1 are:

$$\bar{R}^{(1)}_{\omega} = \left[r_{\omega} + jr\theta_{\omega}\right] e^{j\theta} \tag{B.2}$$

$$\bar{R}^{(2)}_{\omega\omega} = \left[(r_{\omega\omega} - r\theta^2_{\omega}) + j(r\theta_{\omega\omega} + 2r_{\omega}\theta_{\omega})\right] e^{j\theta} \tag{B.3}$$

$$\begin{aligned}\bar{R}^{(3)}_{\omega\omega\omega} = &\left[(r_{\omega\omega\omega} - 3r_{\omega}\theta^2_{\omega} - 3r\theta_{\omega}\theta_{\omega\omega})\right. \\ &\left. + j(r\theta_{\omega\omega\omega} + 3r_{\omega}\theta_{\omega\omega} + 3r_{\omega\omega}\theta_{\omega} - r\theta^3_{\omega})\right] e^{j\theta}.\end{aligned} \tag{B.4}$$

The quantities on the left are the derivatives calculated in Equations 3.13-3.15. It is the phase derivatives (θ_{ω}, $\theta_{\omega\omega}$, and $\theta_{\omega\omega\omega}$) which are generally of interest, however.

To find the phase derivatives, one simply has to solve Equations B.2-B.4. To do so, the real and imaginary components of $\bar{R}^{(m)}e^{-j\theta}$ must be matched

to the real and imaginary components inside the brackets on the right side of Equations B.2-B.4. The real and imaginary components are defined as the parallel ($\hat{\rho}$) and orthogonal ($\hat{\theta}$) projections of the change in reflection onto the unperturbed reflection:

$$\bar{R}^{(1)}_{\omega,\hat{\rho}} = Re\{\bar{R}^{(1)}_{\omega} e^{-j\theta}\} \tag{B.5}$$

$$\bar{R}^{(1)}_{\omega,\hat{\theta}} = Im\{\bar{R}^{(1)}_{\omega} e^{-j\theta}\}. \tag{B.6}$$

Finally, the phase derivatives may be expressed as:

$$r_{\omega} = \bar{R}^{(1)}_{\omega,\hat{r}} \tag{B.7}$$

$$\theta_{\omega} = \frac{1}{r}\bar{R}^{(1)}_{\omega,\hat{\theta}} \tag{B.8}$$

$$r_{\omega\omega} = \bar{R}^{(2)}_{\hat{r}} + r\theta^2_{\omega} \tag{B.9}$$

$$\theta_{\omega\omega} = \frac{1}{r}\left(\bar{R}^{(2)}_{\omega\omega,\hat{\theta}} - 2r_{\omega}\theta_{\omega}\right) \tag{B.10}$$

$$\theta_{\omega\omega\omega} = \frac{1}{r}\left(\bar{R}^{(3)}_{\omega\omega\omega,\hat{\theta}} - 3r_{\omega}\theta_{\omega\omega} - 3r_{\omega\omega}\theta_{\omega} + r\theta^3_{\omega}\right). \tag{B.11}$$

B.2 Two-Dimensional Phase Derivatives

The calculation of group velocity dispersion and third-order dispersion of waveguides in Chapter 4 requires partial derivatives with respect to frequency. Here the derivation proceeds the same as before, but the magnitude and phase are now functions of both frequency and the transverse propagation constant, k_z. That is,

$$\bar{R}(k_z,\omega) = r(k_z,\omega)e^{j\theta(k_z,\omega)}. \tag{B.12}$$

The partial derivatives are:

$$\bar{R}^{(2)}_{k_z\omega} = \left[(r_{k_z\omega} - r\theta_{k_z}\theta_\omega) + j(r\theta_{k_x\omega} + r_\omega\theta_{k_z} + r_{k_z}\theta_\omega)\right] e^{j\theta} \tag{B.13}$$

$$\begin{aligned}\bar{R}^{(3)}_{k_z\omega\omega} &= \big[r_{k_z\omega\omega} - r_{k_z}\theta^2_\omega - 2r_\omega\theta_{k_z}\theta_\omega - r\theta_{k_z}\theta_{\omega\omega} - 2r\theta_\omega\theta_{k_z\omega} \\ &\quad + j(2r_{k_z\omega}\theta_\omega + r_{\omega\omega}\theta_{k_z} + r_{k_z}\theta_{\omega\omega} + 2r_\omega\theta_{k_z\omega} + r\theta_{k_z\omega\omega} - r\theta_{k_z}\theta^2_\omega)\big] e^{j\theta}\end{aligned} \tag{B.14}$$

$$\begin{aligned}\bar{R}^{(3)}_{k_zk_z\omega} &= \big[r_{k_zk_z\omega} - r_\omega\theta^2_{k_z} - 2r_{k_z}\theta_{k_z}\theta_\omega - r\theta_\omega\theta_{k_zk_z} - 2r\theta_{k_z}\theta_{k_z\omega} \\ &\quad + j(2r_{k_z\omega}\theta_{k_z} + r_{k_zk_z}\theta_\omega + r_\omega\theta_{k_zk_z} + 2r_{k_z}\theta_{k_z\omega} + r\theta_{k_zk_z\omega} - r\theta^2_{k_z}\theta_\omega)\big] e^{j\theta}.\end{aligned} \tag{B.15}$$

Solving Equations B.13-B.15 for the partial derivatives with respect to phase, we find:

$$r_{k_x\omega} = \bar{R}^{(2)}_{k_z\omega,\hat{r}} + r\theta_{k_z}\theta_\omega \tag{B.16}$$

$$\theta_{k_z\omega} = \frac{1}{r}\left(\bar{R}^{(2)}_{k_z\omega,\hat{\theta}} - r_\omega\theta_{k_z} - r_{k_z}\theta_\omega\right) \tag{B.17}$$

$$\theta_{k_z\omega\omega} = \frac{1}{r}\left(\bar{R}^{(3)}_{k_z\omega\omega,\hat{\theta}} - 2r_{k_z\omega}\theta_\omega - r_{\omega\omega}\theta_{k_z} - r_{k_z}\theta_{\omega\omega} - 2r_\omega\theta_{k_z\omega} + r\theta_{k_z}\theta^2_\omega\right) \tag{B.18}$$

$$\theta_{k_zk_z\omega} = \frac{1}{r}\left(\bar{R}^{(3)}_{k_zk_z\omega,\hat{\theta}} - 2r_{k_z\omega}\theta_{k_z} - r_{k_zk_z}\theta_\omega - r_\omega\theta_{k_zk_z} - 2r_{k_z}\theta_{k_z\omega} + r\theta^2_{k_z}\theta_\omega\right). \tag{B.19}$$

APPENDIX C

AUXILIARY POTENTIALS

Although Maxwell's equations may always be solved directly in terms of the $\vec{E}$, $\vec{B}$, $\vec{D}$, and $\vec{H}$ fields, solving with auxiliary vector potentials can be more convenient. From an educated guess of the form of the solutions, vector potentials allow the individual components of $\vec{E}$, $\vec{B}$, $\vec{D}$, and $\vec{H}$ to be rapidly calculated. In this appendix, the electric and magnetic fields will be derived from the auxiliary vector potentials.

C.1 Vector Calculus Primer

The derivation of solutions to Maxwell's equations using vector potentials uses theorems from vector calculus which are frequently unfamiliar. Some important theorems are therefore presented here for later use in the actual derivations. For illustrative purposes, let $\vec{C}$ and $\vec{G}$ be arbitrary vector fields and ρ be a scalar field.

A vector field whose divergence is zero ($\nabla\vec{C} = 0$) is defined to be solenoidal. If a vector field $\vec{C}$ is solenoidal, then there exists another vector field $\vec{G}$ such that [43]:

$$\vec{C} = \nabla \times \vec{G}. \tag{C.1}$$

A vector field whose curl is zero ($\nabla \times \vec{C} = 0$) is defined to be irrotational. If a vector field $\vec{C}$ is irrotational, then there exists a scalar field ρ such that [43]:

$$\vec{C} = \nabla\rho. \tag{C.2}$$

The fundamental theorem of vector calculus states that any smooth vector field may be written as the sum of an irrotational and a solenoidal component.

$$\vec{C} = \vec{C}_{sol} + \vec{C}_{irr}. \tag{C.3}$$

One final useful identity for simplifying the curl of the curl of a vector is:

$$\nabla \times (\nabla \times \vec{C}) = \nabla(\nabla \vec{C}) - \nabla^2 \vec{C}. \tag{C.4}$$

Note the parentheses because the order in which the curls are taken is important.

For ease of reference, Equations 2.5 through 2.8 are reproduced below:[1]

$$\nabla \times \vec{E} = j\omega \vec{B} \tag{2.5}$$
$$\nabla \times \vec{H} = -j\omega \vec{D} + \vec{J} \tag{2.6}$$
$$\nabla \cdot \vec{D} = \rho \tag{2.7}$$
$$\nabla \cdot \vec{B} = 0. \tag{2.8}$$

C.2 TM Modes

Equation 2.8 shows that $\vec{B}$ is a solenoidal vector field. We may, therefore, define a vector $\vec{A}$ such that:

$$\vec{B}_A = \nabla \times \vec{A} \tag{C.5}$$

or because most materials are non-magnetic at optical frequencies,

$$\vec{H}_A = \frac{1}{\mu} \nabla \times \vec{A}. \tag{C.6}$$

The vector $\vec{A}$ is known as the magnetic vector potential. It is not a measurable field, but only a mathematical convenience. Inserting this into Equation 2.5, we obtain:

$$\nabla \times \vec{E}_A = j\omega \nabla \times \vec{A} \tag{C.7}$$

or

$$\nabla \times \left(\vec{E}_A - j\omega \vec{A} \right) = 0. \tag{C.8}$$

[1] Note the use of the $e^{-j\omega t}$ convention for time dependence.

The vector on the left side of the previous equation is irrotational, so it may be defined as the gradient of a scalar.

$$-\nabla\varphi_A \equiv \vec{E}_A - j\omega\vec{A} \tag{C.9}$$

or, solving for $\vec{E}_A$:

$$\vec{E}_A = j\omega\vec{A} - \nabla\varphi_A. \tag{C.10}$$

Now, we insert Equations C.6 and C.9 into Equation 2.6 to obtain:

$$\nabla \times (\frac{1}{\mu}\nabla \times \vec{A}) = -j\omega\epsilon(j\omega\vec{A} - \nabla\varphi_A). \tag{C.11}$$

By use of vector calculus identities, we can simplify the left side of the above equation to

$$\nabla(\nabla \cdot \vec{A}) - \nabla^2\vec{A} = \omega^2\mu\epsilon\vec{A} + j\omega\mu\epsilon\nabla\varphi_A. \tag{C.12}$$

The curl of $\vec{A}$ is defined by Equation C.5, but the divergence is unspecified. It is easy to see that any divergenceless vector may be added to $\vec{A}$ without affecting the value of $\vec{B}$ (which is a physically measurable field and therefore must be uniquely defined). We are free to choose any convenient value for the divergence of $\vec{A}$. This is known as setting the gauge.

In Cartesian or cylindrical coordinates, it is convenient to choose

$$\nabla \cdot \vec{A} = j\omega\mu\epsilon\varphi_A \tag{C.13}$$

to simplify Equation C.12, which becomes

$$\nabla^2\vec{A} = \frac{\omega^2\epsilon_r}{c^2}\vec{A}. \tag{C.14}$$

This equation is convenient to solve in rectangular coordinates because $\nabla^2(\hat{x}A_x) = \hat{x}\nabla^2A_x$, $\nabla^2(\hat{y}A_y) = \hat{y}\nabla^2A_y$, and $\nabla^2(\hat{z}A_z) = \hat{z}\nabla^2A_z$. Similarly, in cylindrical coordinates, $\nabla^2(\hat{z}A_z) = \hat{z}\nabla^2A_z$. However, the same pattern does not follow for the $\hat{r}$ and $\hat{\phi}$ components.

In spherical coordinates, there is no coordinate κ such that $\nabla^2(\hat{\kappa}F_\kappa) = \hat{\kappa}\nabla^2F_\kappa$, where $\kappa \in \{r, \phi, \theta\}$. To find a simple wave equation, it will be useful to choose a different gauge. Choosing $\vec{A} = \hat{r}A_r$ and expanding Equation

C.12, we get

$$\frac{1}{\sin\theta}\left[-\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{r}\frac{\partial F_r}{\partial\theta}\right)-\frac{\partial}{\partial\phi}\left(\frac{1}{r\sin\theta}\frac{\partial F_r}{\partial\phi}\right)\right]=\omega^2\mu\epsilon F_r-j\omega\mu\epsilon\frac{\partial\varphi_A}{\partial r} \tag{C.15}$$

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(\frac{\partial F_r}{\partial r}\right)=-\frac{j\omega\mu\epsilon}{r\sin\theta}\frac{\partial\varphi_A}{\partial\phi} \tag{C.16}$$

$$\frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{\partial F_r}{\partial r}\right)=-\frac{j\omega\mu\epsilon}{r}\frac{\partial\varphi_A}{\partial\theta}, \tag{C.17}$$

where Equations C.15, C.16, and C.17 are the $\hat{r}$, $\hat{\phi}$, and $\hat{\theta}$ components respectively. Choosing

$$\nabla\cdot\vec{A}_r=-j\omega\mu\epsilon\varphi_F \tag{C.18}$$

will satisfy Equations C.16-C.17. The $\hat{r}$ equation then becomes

$$\left(\nabla^2+k^2\right)\frac{F_r}{r}=0. \tag{C.19}$$

This can be solved by defining $\psi_A=\frac{A_r}{r}$. In the older literature, ψ_A is often referred to as the Debye potential [44].

After Equation C.14 or C.19 has been solved, the electric and magnetic fields corresponding to the TM fields components can be found from:

$$\vec{E}_A=j\omega\vec{A}-\frac{1}{j\omega\mu\epsilon}\nabla(\nabla\cdot\vec{A}) \tag{C.20}$$

$$\vec{H}_A=\frac{1}{\mu}\nabla\times\vec{A}. \tag{C.21}$$

From Equation C.21, it is obvious that $\vec{H}_A$ is perpendicular to $\vec{A}$. Because the magnetic field is perpendicular to the direction of $\vec{A}$, these waves are known as the transverse magnetic, or TM, waves.

C.3 TE Modes

The line of reasoning to find TE modes is the same as that for TM modes. In a medium with no charge, Equation 2.7 shows that $\vec{D}$ is a solenoidal field.

Therefore, there exists a vector $\vec{F}$ such that

$$\vec{D}_F = -\nabla \times \vec{F} \tag{C.22}$$

which, in a uniform medium, becomes:

$$\vec{E}_F = -\frac{1}{\epsilon}\nabla \times \vec{F}. \tag{C.23}$$

Inserting Equation C.22 into Equation 2.6, we obtain:

$$\nabla \times \vec{H_F} = -j\omega(-\nabla \times \vec{F}) \tag{C.24}$$

or

$$\nabla \times (\vec{H}_F - j\omega\vec{F}) = 0. \tag{C.25}$$

The vector on the left side of the previous equation is irrotational, so it may be defined as the gradient of a scalar function:

$$\nabla\varphi_F \equiv \vec{H}_F - j\omega\vec{F}. \tag{C.26}$$

Solving for H_F, we find:

$$\vec{H}_F = \nabla\varphi_F + j\omega\vec{F}. \tag{C.27}$$

Inserting Equations C.23 and C.27 into the left and right sides of Equation 2.5, respectively, a wave equation for the field $\vec{F}$ is obtained:

$$\nabla \times (-\frac{1}{\epsilon}\nabla \times \vec{F}) = j\omega\mu(j\omega\vec{F} + \nabla\varphi_F). \tag{C.28}$$

In a uniform material, $\frac{1}{\epsilon}$ may be factored out of the curl equation and the result simplified to

$$(\nabla \cdot \vec{F}) - \nabla^2\vec{F} = \omega^2\mu\epsilon\vec{F} - j\omega\mu\epsilon\nabla\varphi_F. \tag{C.29}$$

Once again, we are free to choose the gauge. In Cartesian and cylindrical coordinates, it is convenient to set

$$\nabla \cdot \vec{F} = -j\omega\mu\epsilon\varphi_F \tag{C.30}$$

so that Equation C.29 becomes

$$\nabla^2 \vec{F} = \frac{\omega^2 \epsilon_r}{c^2} \vec{F}. \tag{C.31}$$

In spherical coordinates, there is no coordinate κ such that $\nabla^2(\hat{\kappa} F_\kappa) = \hat{\kappa} \nabla^2 F_\kappa$, where $\kappa \in \{r, \phi, \theta\}$. To find a simple wave equation, it will be useful to choose a different gauge. Choosing $\vec{F} = \hat{r} F_r$ and expanding Equation C.29, we find the $\hat{r}$, $\hat{\phi}$, and $\hat{\theta}$ are, respectively

$$\frac{1}{\sin\theta}\left[-\frac{\partial}{\partial\theta}\left(\frac{\sin\theta}{r}\frac{\partial F_r}{\partial\theta}\right) - \frac{\partial}{\partial\phi}\left(\frac{1}{r\sin\theta}\frac{\partial F_r}{\partial\phi}\right)\right] = \omega^2\mu\epsilon F_r - j\omega\mu\epsilon\frac{\partial\varphi_F}{\partial r} \tag{C.32}$$

$$\frac{1}{r\sin\theta}\frac{\partial}{\partial\phi}\left(\frac{\partial F_r}{\partial r}\right) = -\frac{j\omega\mu\epsilon}{r\sin\theta}\frac{\partial\varphi_F}{\partial\phi} \tag{C.33}$$

$$\frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{\partial F_r}{\partial r}\right) = -\frac{j\omega\mu\epsilon}{r}\frac{\partial\varphi_F}{\partial\theta}. \tag{C.34}$$

Choosing

$$\nabla \cdot \vec{F}_r = -j\omega\mu\epsilon\varphi_F \tag{C.35}$$

will satisfy Equations C.33-C.34. The $\hat{r}$ equation then becomes

$$\left(\nabla^2 + k^2\right)\frac{F_r}{r} = 0. \tag{C.36}$$

This can be solved by defining $\psi_F = \frac{F_r}{r}$. In the older literature, ψ_F is referred to as the Debye potential [44].

Once either Equation C.31 or C.36, has been solved the electric and magnetic fields may be calculated via the expressions:

$$\vec{E}_F = -\frac{1}{\epsilon}\nabla \times \vec{F} \tag{C.37}$$

$$\vec{H}_F = j\omega\vec{F} + \nabla\varphi_F = j\omega\vec{F} - \frac{1}{j\omega\mu\epsilon}\nabla(\nabla \cdot \vec{F}). \tag{C.38}$$

From the Equation C.37, it is readily apparent that $\vec{E}_F$ is perpendicular to $\vec{F}$. The electric vector potential, therefore, generates transverse electric, or TE, waves.

APPENDIX D

DELVES ALGORITHM

The Delves algorithm [32] is a computational method used to find complex roots of an analytic function within a bounded region. It was used by Chilwell and Hodgkinson [8] in 1984 to locate the lossy modes a of a waveguide. This appendix will explain the basic aspects of the Delves algorithm.

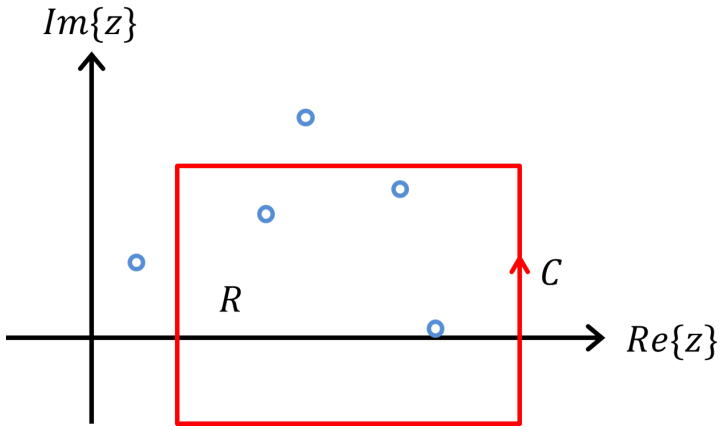

Figure D.1: Layout of the Delves algorithm. The zeros outside of the curve C will not be located by the algorithm.

Let $f(z)$ be an analytic function whose zeros are to be determined in region R enclosed within a counterclockwise curve C, as shown in Figure D.1. Furthermore, it is assumed that no poles of f occur within R. The Delves algorithm comprises the following steps:

1. Calculate the number of zeros in R using Cauchy's argument principle.
   (a) If necessary, divide the region into smaller sub-regions so that each subregion has only a limited number of zeros.
   (b) If the curve C (or any sub-curves) passes too close to a root, move it outwards for numerical stability.

2. Use the argument principle to calculate the power sums of the roots in each region.

3. Use the Newton-Girard formulas to construct in each subregion a polynomial which has the same roots as f and solve them with a polynomial root finding algorithm.

4. If the boundary curve was moved in step 1b, remove any excess zeros which are outside the original boundary.

The first step of the algorithm is to find the number of zeros in R. The argument principle [45] states that

$$N - P = \frac{1}{j2\pi} \oint_C \frac{f'(z)}{f(z)} dz, \tag{D.1}$$

where N is the number of zeros and P is the number of poles. The integral may be evaluated numerically to determine the number of zeros enclosed in C. Delves and Lyness note [32] that this integral does not need to be carried out to high precision because the result is known to be an integer.

In the second step of the algorithm, the region R is divided into subregions, R_i, and both the internal and external boundaries are adjusted to avoid numerical difficulties. In step 3 of the algorithm, a polynomial will be constructed which possesses the same roots of $f(z)$ within R (or the subregions R_i). Numerical techniques will then be used to find the roots of this polynomial (and hence the roots of $f(z)$). Algorithms to find the roots of a polynomial become numerically unstable if the number of roots becomes large. The region R is, therefore, recursively sub-divided until each region R_i contains only a limited number of zeros (generally less than 5). A simple example of this division is shown in Figure D.2.

Additionally, the boundaries C_i among the subregions R_i may need to be moved to increase the accuracy of the calculation. If the curve C passes too close to a zero, the integrand in Equation D.1 may become very large, potentially causing a loss of precision when the integral is evaluated numerically with floating-point arithmetic. Because at this point in the algorithm the zeros are unknown, nearness to a zero is inferred from the value of $f'(z)/f(z)$ becoming large. Delves and Lyness suggest several methods for how to move the boundaries [32]. Many other methods may be invented. During this

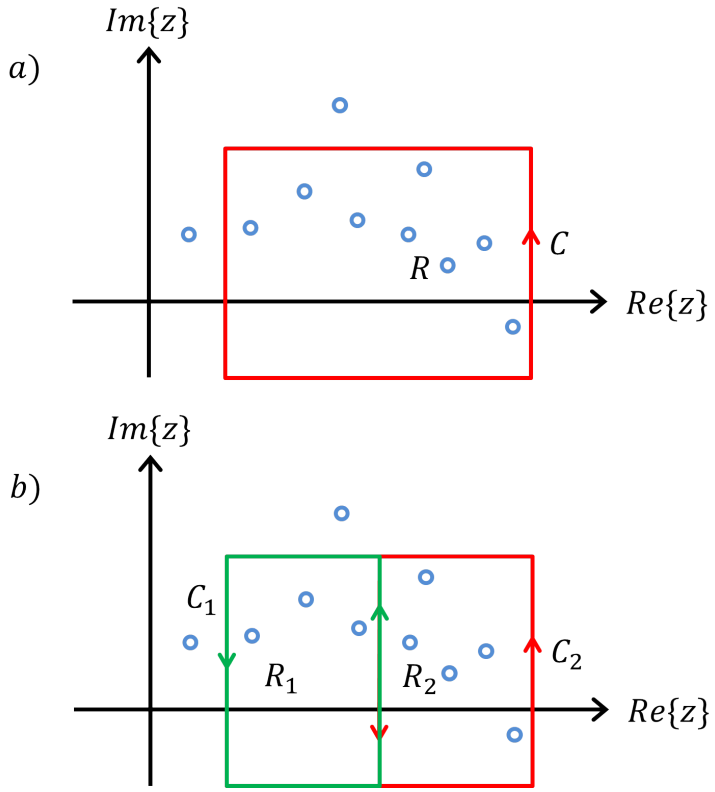

Figure D.2: Region R has too many zeros, so it is divided into subregions R_1 and R_2. The boundary curves C_1 and C_2 remain oriented counter-clockwise.

process, subregions may become overlapping, resulting in double-counting of roots and the inclusion of roots outside the original region R. This is remedied in the final step of the algorithm.

The third step of the algorithm is to construct and solve a polynomial with the same zeros as $f(z)$ in each subregion R_i. The polynomial is defined as

$$p(z) = \prod_{q=1}^{N} (z - z_q) = \sum_{q=0}^{N} (-1)^{N+q} p_{N-q} z^q, \tag{D.2}$$

where z_q's are the roots of the polynomial, N is the number of roots inside R_i, and p_q are the coefficients of the corresponding polynomial. For each subregion, we must now calculate the power sums of the roots. A power sum S_m is defined as the sum of the roots to the m^{th} power:

$$S_m^{(i)} \equiv \sum_k z_k^m. \tag{D.3}$$

Conveniently, the argument principle provides a method to calculate these sums numerically:

$$S_m = \frac{1}{j2\pi} \oint_C z^m \frac{f'(z)}{f(z)} dz. \tag{D.4}$$

The Newton-Girard formulas relate the coefficients of a polynomial to the power sums of the roots of that polynomial. The coefficients of the polynomial are:

$$p_0 = 1 \tag{D.5}$$

$$p_1 = S_1 \tag{D.6}$$

$$p_2 = \frac{1}{2}(p_1 S_1 - S_2) \tag{D.7}$$

$$p_3 = \frac{1}{3}(p_2 S_1 - p_1 S_1 + S_3) \tag{D.8}$$

$$p_4 = \frac{1}{4}(p_3 S_1 - p_2 S_2 + p_1 S_3 - S_4). \tag{D.9}$$

Finally, the roots of the polynomial may be found with a polynomial root finder.

APPENDIX E

BESSEL FUNCTIONS

When working cylindrical and spherical coordinates, equations will often arise involving Bessel functions, modified Bessel functions, spherical Bessel functions, Riccati-Bessel functions, and their derivatives. Many software packages (such as MATLAB) do not have built-in libraries to evaluate these functions. This appendix provides useful identities which allow the calculation of any of these functions requiring only access to a mathematical package which can compute the ordinary Bessel functions $J_m(x)$ and $Y_m(x)$. The identities for the first derivative of the Bessel functions and the definition for the Riccati-Bessel functions have been collected from the Digital Library of Mathematical Functions [46, 47].

E.1 Derivatives of Cylindrical Bessel Functions

Let B_m represent any linear combination of J_m, Y_m, $H_m^{(1)}$, or $H_m^{(2)}$. Then, the derivative of B_m is:

$$B'_m(x) = \begin{cases} -B_1(x) & m = 0 \\ \frac{1}{2}\left[B_{m-1}(x) - B_{m+1}(x)\right] & m \neq 0. \end{cases} \tag{E.1}$$

The second derivative of B_m is:

$$B''_m(x) = \begin{cases} \frac{1}{2}\left[B_2(x) - B_0(x)\right] & m = 0 \\ \frac{1}{4}\left[B_3(x) - 3B_1(x)\right] & m = 1. \\ \frac{1}{4}\left[B_{m-2}(x) - 2B_m(x) + B_{m+2}(x)\right] & m \neq 0, 1. \end{cases} \tag{E.2}$$

Though only listed up to second order, derivatives of any order may be found by using Equation E.1 recursively.

E.2 Derivatives of Modified Bessel Functions

The modified Bessel functions, $I_m(x)$ and $K_m(x)$, are obtained from the cylindrical Bessel functions, $J_m(x)$ and $Y_m(x)$, by substituting jx for x. In this section, let $B_m(x)$ represent either $I_m(x)$ or $(-1)^m K_m(x)$. Then, the first derivatives of the modified Bessel functions are:

$$B'_m(x) = \begin{cases} B'_0(x) = -B_1(x) & m = 0 \\ \frac{1}{2}\left(B_{m-1}(x) - B_{m+1}(x)\right) & m \neq 0. \end{cases} \tag{E.3}$$

The second derivatives are:

$$B''_m(x) = \begin{cases} \frac{1}{2}\left(B_2(x) - B_0(x)\right) & m = 0 \\ \frac{1}{4}\left(B_3(x) - 3B_1(x)\right) & m = 1 \\ \frac{1}{4}\left(B_{m-2}(x) - 2B_m(x) + B_{m+2}(x)\right) & m \neq 0, 1. \end{cases} \tag{E.4}$$

Similar to cylindrical Bessel functions, derivatives of modified Bessel functions may be found to any order by using Equation E.3 recursively.

E.3 Riccati-Bessel Functions and Their Derivatives

The Riccati-Bessel functions occur in the solutions to wave equations in spherical coordinates. In terms of cylindrical Bessel functions, they are defined as:

$$\hat{J}_m(x) = \sqrt{\frac{\pi x}{2}} J_{m+\frac{1}{2}}(x) \tag{E.5}$$

$$\hat{Y}_m(x) = -\sqrt{\frac{\pi x}{2}} Y_{m+\frac{1}{2}}(x) \tag{E.6}$$

$$\hat{H}^{(1)}_m(x) = \sqrt{\frac{\pi x}{2}} H^{(1)}_{m+\frac{1}{2}}(x) \tag{E.7}$$

$$\hat{H}^{(2)}_m(x) = \sqrt{\frac{\pi x}{2}} H^{(2)}_{m+\frac{1}{2}}(x). \tag{E.8}$$

The first derivatives of the Riccati-Bessel functions can likewise be expressed in terms of Bessel functions.

$$\hat{J}'_m(x) = \sqrt{\frac{\pi x}{2}} J'_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{8x}} J_{m+\frac{1}{2}}(x) \tag{E.9}$$

$$\hat{Y}'_m(x) = -\sqrt{\frac{\pi x}{2}} Y'_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{8x}} Y_{m+\frac{1}{2}}(x) \tag{E.10}$$

$$\hat{H}^{(1)\prime}_m(x) = \sqrt{\frac{\pi x}{2}} H^{(1)\prime}_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{8x}} H^{(1)}_{m+\frac{1}{2}}(x) \tag{E.11}$$

$$\hat{H}^{(2)\prime}_m(x) = \sqrt{\frac{\pi x}{2}} H^{(2)\prime}_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{8x}} H^{(2)}_{m+\frac{1}{2}}(x). \tag{E.12}$$

Furthermore, the second derivatives are:

$$\hat{J}''_m(x) = \sqrt{\frac{\pi x}{2}} J''_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{2x}} J'_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{32x^3}} J_{m+\frac{1}{2}}(x) \tag{E.13}$$

$$\hat{Y}''_m(x) = -\sqrt{\frac{\pi x}{2}} Y''_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{2x}} Y'_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{32x^3}} Y_{m+\frac{1}{2}}(x) \tag{E.14}$$

$$\hat{H}^{(1)\prime\prime}_m(x) = \sqrt{\frac{\pi x}{2}} H^{(1)\prime\prime}_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{2x}} H^{(1)\prime}_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{32x^3}} H^{(1)}_{m+\frac{1}{2}}(x) \tag{E.15}$$

$$\hat{H}^{(2)\prime\prime}_m(x) = \sqrt{\frac{\pi x}{2}} H^{(2)\prime\prime}_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{2x}} H^{(2)\prime}_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{32x^3}} H^{(2)}_{m+\frac{1}{2}}(x). \tag{E.16}$$

E.4 Spherical Bessel Functions and Their Derivatives

Spherical Bessel functions also occur in the solutions to wave equations in spherical coordinates. In terms of cylindrical Bessel functions and Riccati-Bessel functions, they are defined as:

$$j_m(x) = \frac{1}{x}\hat{J}_m(x) = \sqrt{\frac{\pi}{2x}} J_{m+\frac{1}{2}}(x) \tag{E.17}$$

$$y_m(x) = \frac{1}{x}\hat{Y}_m(x) = -\sqrt{\frac{\pi}{2x}} Y_{m+\frac{1}{2}}(x) \tag{E.18}$$

$$h^{(1)}_m(x) = \frac{1}{x}\hat{H}^{(1)}_m(x) = \sqrt{\frac{\pi}{2x}} H^{(1)}_{m+\frac{1}{2}}(x) \tag{E.19}$$

$$h^{(2)}_m(x) = \frac{1}{x}\hat{H}^{(2)}_m(x) = \sqrt{\frac{\pi}{2x}} H^{(2)}_{m+\frac{1}{2}}(x). \tag{E.20}$$

The first derivatives of the Riccati-Bessel functions can likewise be expressed in terms of Bessel functions:

$$j'_m(x) = \sqrt{\frac{\pi}{2x}} J'_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{8x^3}} J_{m+\frac{1}{2}}(x) \tag{E.21}$$

$$y'_m(x) = -\sqrt{\frac{\pi}{2x}} Y'_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{8x^3}} Y_{m+\frac{1}{2}}(x) \tag{E.22}$$

$$h^{(1)\prime}_m(x) = \sqrt{\frac{\pi}{2x}} H^{(1)\prime}_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{8x^3}} H^{(1)}_{m+\frac{1}{2}}(x) \tag{E.23}$$

$$h^{(2)\prime}_m(x) = \sqrt{\frac{\pi}{2x}} H^{(2)\prime}_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{8x^3}} H^{(2)}_{m+\frac{1}{2}}(x). \tag{E.24}$$

Furthermore, the second derivatives are:

$$j''_m(x) = \sqrt{\frac{\pi}{2x}} J''_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{2x^3}} J'_{m+\frac{1}{2}}(x) + \sqrt{\frac{9\pi}{32x^5}} J_{m+\frac{1}{2}}(x) \tag{E.25}$$

$$y''_m(x) = -\sqrt{\frac{\pi}{2x}} Y''_{m+\frac{1}{2}}(x) + \sqrt{\frac{\pi}{2x^3}} Y'_{m+\frac{1}{2}}(x) - \sqrt{\frac{9\pi}{32x^5}} Y_{m+\frac{1}{2}}(x) \tag{E.26}$$

$$h^{(1)\prime\prime}_m(x) = \sqrt{\frac{\pi}{2x}} H^{(1)\prime\prime}_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{2x^3}} H^{(1)\prime}_{m+\frac{1}{2}}(x) + \sqrt{\frac{9\pi}{32x^5}} H^{(1)}_{m+\frac{1}{2}}(x) \tag{E.27}$$

$$h^{(2)\prime\prime}_m(x) = \sqrt{\frac{\pi}{2x}} H^{(2)\prime\prime}_{m+\frac{1}{2}}(x) - \sqrt{\frac{\pi}{2x^3}} H^{(2)\prime}_{m+\frac{1}{2}}(x) + \sqrt{\frac{9\pi}{32x^5}} H^{(2)}_{m+\frac{1}{2}}(x). \tag{E.28}$$

REFERENCES

[1] H. A. Macleod, *Thin Film Optical Filters*, 4th ed. Boca Raton, FL: CRC Press, 2010.

[2] H. Soda, K. Iga, C. Kitahara, and Y. Suematsu, "GaInAsP/InP surface emitting injection lasers," *Japanese Journal of Applied Physics*, vol. 18, no. 12, p. 2329, 1979.

[3] K. Iga, F. Koyama, and S. Kinoshita, "Surface emitting semiconductor lasers," *IEEE Journal of Quantum Electronics*, vol. 24, no. 9, pp. 1845–1855, Sept 1988.

[4] V. Braginsky, M. Gorodetsky, and V. Ilchenko, "Quality-factor and nonlinear properties of optical whispering-gallery modes," *Physics Letters A*, vol. 137, no. 78, pp. 393 – 397, 1989.

[5] B. Prade, J. Y. Vinet, and A. Mysyrowicz, "Guided optical waves in planar heterostructures with negative dielectric constant," *Phys. Rev. B*, vol. 44, pp. 13 556–13 572, Dec 1991.

[6] M. G. Moharam, D. A. Pommet, E. B. Grann, and T. K. Gaylord, "Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach," *J. Opt. Soc. Am. A*, vol. 12, no. 5, pp. 1077–1086, May 1995.

[7] D. M. Pai and K. A. Awada, "Analysis of dielectric gratings of arbitrary profiles and thicknesses," *J. Opt. Soc. Am. A*, vol. 8, no. 5, pp. 755–762, May 1991.

[8] J. Chilwell and I. Hodgkinson, "Thin-films field-transfer matrix theory of planar multilayer waveguides and reflection from prism-loaded waveguides," *J. Opt. Soc. Am. A*, vol. 1, no. 7, pp. 742–753, Jul 1984.

[9] C. Chen, P. Berini, D. Feng, S. Tanev, and V. Tzolov, "Efficient and accurate numerical analysis of multilayer planar optical waveguides in lossy anisotropic media," *Optics Express*, vol. 7, no. 8, pp. 260–272, 2000.

[10] J. P. Boyd, *Chebyshev and Fourier Spectral Methods*, 2nd ed. Mineola, NY: Dover, 2001.

[11] G. B. Airy, "On the phenomena of Newton's rings when formed between two transparent substances of different refractive powers," *The London and Edinburgh Philisophical Magazine and Journal of Science*, vol. 2, pp. 20–30, Jan 1833.

[12] W. L. Stutzman and G. A. Thiele, *Antenna Theory and Design*, 3rd ed. Hoboken, NJ: Wiley, 2013.

[13] P. Yeh, *Optical Waves in Layered Media*. New York, NY: Wiley, 1988.

[14] R. D. Radcliff, C. Balanis et al., "Modified propagation constants for nonuniform plane wave transmission through conducting media," *IEEE Transactions on Geoscience and Remote Sensing*, no. 3, pp. 408–411, 1982.

[15] A. Stingl, C. Spielmann, F. Krausz, and R. Szipöcs, "Generation of 11-fs pulses from a Ti:sapphire laser without the use of prisms," *Opt. Lett.*, vol. 19, no. 3, pp. 204–206, Feb 1994.

[16] R. Paschotta, G. J. Spühler, D. H. Sutter, N. Matuschek, U. Keller, M. Moser, R. Hövel, V. Scheuer, G. Angelow, and T. Tschudi, "Double-chirped semiconductor mirror for dispersion compensation in femtosecond lasers," *Applied Physics Letters*, vol. 75, no. 15, pp. 2166–2168, 1999.

[17] F. X. Kärtner, U. Morgner, R. Ell, T. Schibli, J. G. Fujimoto, E. P. Ippen, V. Scheuer, G. Angelow, and T. Tschudi, "Ultrabroadband double-chirped mirror pairs for generation of octave spectra," *J. Opt. Soc. Am. B*, vol. 18, no. 6, pp. 882–885, Jun 2001.

[18] N. Matuschek, F. Kartner, and U. Keller, "Analytical design of double-chirped mirrors with custom-tailored dispersion characteristics," *IEEE Journal of Quantum Electronics*, vol. 35, no. 2, pp. 129–137, 1999.

[19] A. V. Tikhonravov and M. K. Trubetskov, "Modern design tools and a new paradigm in optical coating design," *Appl. Opt.*, vol. 51, no. 30, pp. 7319–7332, Oct 2012.

[20] A. Yariv and P. Yeh, *Photonics*, 6th ed. Oxford University Press, 2007.

[21] N. Matuschek, F. Kartner, and U. Keller, "Exact coupled-mode theories for multilayer interference coatings with arbitrary strong index modulations," *IEEE Journal of Quantum Electronics*, vol. 33, no. 3, pp. 295–302, 1997.

[22] J. R. Birge and F. X. Kärtner, "Efficient analytic computation of dispersion from multilayer structures," *Appl. Opt.*, vol. 45, no. 7, pp. 1478–1483, Mar 2006.

[23] A. W. Crook, "The reflection and transmission of light by any system of parallel isotropic films," *J. Opt. Soc. Am.*, vol. 38, no. 11, pp. 954–963, Nov 1948.

[24] R. Trebino, *Frequency-Resolved Optical Gating: The Measurement of Ultrashort Laser Pulses*. Boston, MA: Kluwer Academic Publishers, 2000.

[25] A. Jennings and J. J. McKeown, *Matrix Computation*, 2nd ed. New York, NY: Wiley, 1992.

[26] V. Pervak, S. Naumov, G. Tempea, V. Yakovlev, F. Krausz, and A. Apolonski, "Synthesis and manufacturing the mirrors for ultrafast optics," *Proc. SPIE*, vol. 5963, pp. 59 631P–1, 2005.

[27] M. Trubetskov, A. Tikhonravov, and V. Pervak, "Time-domain approach for designing dispersive mirrors based on the needle optimization technique. Theory." *Opt. Express*, vol. 16, no. 25, pp. 20 637–20 647, Dec 2008.

[28] R. A. Horn and C. R. Johnson, *Matrix Analysis*, 2nd ed. New York, NY: Cambridge, 2013.

[29] H. C. van de Hulst, *Light Scattering by Small Particles*. New York, NY: Wiley, 1957.

[30] J.-M. Jin, *Theory and Computation of Electromagnetic Fields*, 1st ed. Hoboken, NJ: Wiley, 2010.

[31] R. C. Rumpf, "Improved formulation of scattering matrices for semi-analytical methods that is consistent with convention," *Progress in Electromagnetics Research B*, vol. 35, pp. 241–261, 2011.

[32] L. Delves and J. Lyness, "A numerical method for locating the zeros of an analytic function," *Mathematics of Computation*, vol. 21, no. 100, pp. 543–560, 1967.

[33] C. G. B. Garrett, W. Kaiser, and W. L. Bond, "Stimulated emission into optical whispering modes of spheres," *Phys. Rev.*, vol. 124, pp. 1807–1809, Dec 1961.

[34] P. L. Gourley, K. E. Meissner, T. M. Brennan, B. E. Hammons, and M. F. Gourley, "Surface-emitting semiconductor laser spectroscopy and microscopy for characterizing normal and sickled red blood cells," *Proc. SPIE*, vol. 2387, pp. 148–161, 1995.

[35] K. H. Drexhage, "Structure and properties of laser dyes," in *Dye lasers*. Springer, 1973, pp. 144–193.

[36] H. Kogelnik and T. Li, "Laser beams and resonators," *Appl. Opt.*, vol. 5, no. 10, pp. 1550–1567, Oct 1966.

[37] J. T. Verdeyen, *Laser Electronics*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 1995.

[38] A. E. Siegman, *Lasers*, 1st ed. Sausalito, CA: University Science Books, 1986.

[39] W. T. Silfvast, *Laser Fundamentals*, 2nd ed. Cambridge, UK: Cambridge University Press, 2004.

[40] T. A. Jr., E. B. Bradford, J. W. Vanderhoff, and G. Oster, "Optical properties of uniform particle-size latexes," *J. Opt. Soc. Am.*, vol. 44, no. 8, pp. 603–607, Aug 1954.

[41] N. Denkov, O. Velev, P. Kralchevski, I. Ivanov, H. Yoshimura, and K. Nagayama, "Mechanism of formation of two-dimensional crystals from latex particles on substrates," *Langmuir*, vol. 8, no. 12, pp. 3183–3190, 1992.

[42] N. Denkov, O. Velev, P. Kralchevsky, I. Ivanov, H. Yoshimura, and K. Nagayama, "Two-dimensional crystallization," *Nature*, vol. 361, p. 26, 1993.

[43] W. Kaplan, *Advanced Calculus*, 5th ed. New York, NY: Addison-Wesley, 2003.

[44] A. N. Oraevsky, "Whispering-gallery waves," *Quantum Electronics*, vol. 32, no. 5, pp. 377–400, 2002.

[45] J. W. Brown and R. V. Churchill, *Complex Variables and Applications*, 7th ed. New York, NY: McGraw-Hill, 2004.

[46] F. W. J. Olver, D. W. Lozier, R. F. Boisvert, and C. W. Clark, Eds., *NIST Handbook of Mathematical Functions*. New York, NY: Cambridge University Press, 2010, print companion to [47].

[47] "NIST digital library of mathematical functions," 2014. [Online]. Available: http://dlmf.nist.gov/