

## AN ABSTRACT OF THE THESIS OF

Bradley S. Matson for the degree of Doctor of Philosophy in Physics presented on

31 August, 1998. Title: THEORETICAL ANALYSIS OF OPTICAL MIXING AT INTERFACES.

Abstract approved: Redacted for Privacy

William M. Hetherington III

### ABSTRACT

The art of obtaining interface-specific information via optical mixing has been continuously developed since the advent of the laser in the early sixties. For instance, generating the second harmonic on crystals with inversion symmetry necessarily leads to interface-specific signals by virtue of breaking that symmetry at an interface. In this work, two techniques of interface-specific optical mixing are analyzed theoretically. The first is an exact macroscopic treatment of general optical mixing in reflection geometry. This model is an extension of Bloembergen's model which utilized a surface layer with nonlinear and linear properties atop a semi-infinite bulk with only linear properties. This model was extended by adding a bulk median with the nonlinear properties of the bulk between the surface layer and semi-infinite bulk. This allows for the separate comparison of the surface and bulk contributions to the total signal. Solutions to the model show an overall dependence on the secant of the nonlinear 'reflection' angle in addition to the angular dependencies introduced by the layered structure. Complicated phase-matching dependencies appear in three factors to both the surface and bulk contributions. They include the usual sinc behavior of the phase-mismatch parameter, dependencies on the effective phase differences incurred in the boundary layers, and a dependence on

$(k_t^2 - k_s^2)^{-1}$  where  $k_t$  and  $k_s$  are the wavevectors for the mixed wave and effective source wave caused by the induced polarization. A numerical analysis of these solutions shows that for pure *S*-polarization optical mixing in reflection geometry, surface-specificity can be enhanced using near-grazing incident angles. If the bulk and surface dispersions are small but different, the surface signal can be isolated from the bulk signal. Signals from either *S* or *P*-polarization in a CARS experiment can be selected by average changes in angles of 0.1 degrees. In *P*-polarization cases, it is possible to separate signals by observing the nonlinear 'reflection' at the nonlinear Brewster's angle of the bulk. Since the nonlinear Brewster's angle is different for the two regions, there may be only a surface contribution. A second technique using evanescent fields at a dielectric waveguide interface is analyzed and discussed. Three-beam and four-beam CARS experiments are compared. For non-phase matched conditions the coherence length is small, typically several microns. These are compared with the case where phase-matching is achieved. Despite the shorter interaction length, the phase-matched case generally provides signals two to three times larger. Typically, there is an additional enhancement of 10 to  $10^3$  if dispersions are included. This work shows that for reflection geometry optical mixing the surface contribution can be enhanced relative to the bulk without the experimental difficulty imposed by the use of waveguides.

© Copyright by Bradley S. Matson  
31 August, 1998

All Rights Reserved

# THEORETICAL ANALYSIS OF OPTICAL MIXING AT INTERFACES

by

Bradley S. Matson

A THESIS

submitted to

Oregon State University

in partial fulfillment of  
the requirements for the  
degree of

Doctor of Philosophy

Presented 31 August, 1998

Commencement June 1999



Doctor of Philosophy thesis of Bradley S. Matson presented on 31 August, 1998

APPROVED:

Redacted for Privacy

Major Professor, representing Physics

Redacted for Privacy

Chair of Department of Physics

Redacted for Privacy

Dean of Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Redacted for Privacy

Bradley S. Matson, Author

## ACKNOWLEDGMENTS

I am sincerely grateful to my advisor, Bill Hetherington, for his patience and support over the past seven years. His breadth of knowledge of this field, including both basic phenomena and a much needed connection with experimental reality, have been invaluable. His unending indulgence saw us through many fascinating diversions, yet eventually steering us back to reality. With patience he discussed results and details of the work and with friendship provided a supportive and encouraging atmosphere.

I should also like to thank Jeanette Roberts for many hours of engaging and foresight filled conversations over coffee. These conversations provided many insights and kernels for this work. I am indebted to the solid friendship and support of Ashley Shultz who patiently experienced my triumphs and defeats. Also I owe thanks to my family for their patience, to Naomi Fisher for her constant encouragement, to Tim Taylor for reassurance of my abilities, to my community of friends, and to the whole of the OSU department of physics for the many conversations in the hallways that solidified my confidence in this work. I could not have completed the writing of this thesis without the patience and assistance of Melissa Taylor; my deepest thanks and apologies for my intolerance.

The deepest thanks from my heart is perhaps best expressed in the words of Daniel  
(דניאל ב:כד):

לך אלת אבותי מהודא ומשבח אנה די חכורתא יחבת לי

תושלבע

## TABLE OF CONTENTS

1. INTRODUCTION .....	1
1.1 How This Research Developed.....	4
1.2 Review of Optical Mixing and Definition of Optical Mixing Terms .....	6
1.3 Interface-Specific Optical Mixing (ISOM) .....	11
1.4 ISOM Using Planar Waveguides (WISOM).....	14
1.5 ISOM Using Reflection Geometry (RISOM) .....	18
2. INTERFACE-SPECIFIC OPTICAL MIXING EMPLOYING GUIDED WAVE GEOMETRY .....	23
2.1 Introduction .....	23
2.2 Overview of CARS-Like FWM and WISOM Experiments.....	25
2.3 Previous Experimental WISOM Work.....	37
2.4 Theoretical Development of WISOM .....	38
2.4.1 The Wave Equation for Waveguides .....	40
2.4.2 Review of Solutions to the Wave Equation for Inhomogeneous Waveguides .....	42
2.4.3 Plane Wave Solutions to the Step Index Waveguide.....	45
2.4.4 The Intensity, Phase-Mismatch, and Quantitative Analysis Tools for WISOM .....	50
2.5 Introduction to the Numerical Analysis of WISOM.....	56
2.5.1 Common CARS-Like WISOM Parameters .....	58
2.5.2 How to Interpret the Numerical Results .....	60
2.5.3 Results of the Numerical Analysis .....	67
2.6 Phase-Matching Gaussian Beams in a Waveguide .....	83
2.6.1 Deriving the Phase-Mismatch of WISOM Using Focused Gaussian Beams.....	84

## TABLE OF CONTENTS (Continued)

2.6.2 Results of Phase-Mismatch of WISOM Using Focused Gaussian Beams .....	88
2.7 Couplers for Four-Beam Mode Combinations.....	94
2.7.1 Achromatic Prism Couplers .....	95
2.7.2 Integrated Grating Couplers.....	97
2.7.3 Integrated Nonlinear Grating Couplers.....	100
2.7.4 Ring Gratings .....	101
2.8 Conclusions.....	106
3. REFLECTION GEOMETRY INTERFACE-SPECIFIC OPTICAL MIXING (RISOM) .....	109
3.1 Introduction .....	109
3.2 Microscopic Versus Macroscopic Treatments of Optical Mixing.....	117
3.3 Macroscopic Models Describing Nonlinear Processes at Surfaces .....	122
3.3.1 Previous Work.....	122
3.3.2 Bloembergen's Model in Detail.....	125
3.3.3 Results and Limitations of Bloembergen's Model.....	142
3.4 Developing a Macroscopic Model for RISOM Prediction .....	143
3.4.1 Deriving the Behaviour of the Generated Fields .....	145
3.4.2 Summary and Discussion of the Macroscopic RISOM Model.....	157
3.4.3 Prelude to Numerical Analysis of the RISOM expressions.....	158
3.5 Introduction to the Numerical Analysis of the RISOM Expressions.....	159
3.5.1 Definitions and Common Terms Used in the Analysis.....	166
3.5.2 Parameter Settings.....	167
3.5.3 How to Interpret Analysis Plots .....	168
3.6 Analysis of RISOM in Collinear Geometry .....	169
3.6.1 Collinear General Optical Mixing .....	169
3.6.1.1 S-polarization NLRC for Collinear Optical Mixing in Dispersionless Media .....	170
3.6.1.2 S-polarization NLRC for Collinear $n$ HG in Media with Dispersion.....	173

## TABLE OF CONTENTS (Continued)

3.6.1.3 <i>P</i> -polarization NLRC for Collinear $n$ HG in Dispersionless Media .....	174
3.6.2 <i>P</i> -Wave Collinear Geometry SHG With Dispersion: Further Analysis of the NLBA.....	179
3.6.3 Collinear Geometry SHG With Dispersion: Further Analysis of the NLBA.....	181
3.7 Analysis of RISOM in Non-Collinear Geometry .....	182
3.7.1 Non-Collinear Geometry SHG Without Dispersion .....	182
3.7.1.1 Polar Angle Phase-Mismatched <i>S</i> -Polarization SHG Without Dispersion .....	185
3.7.1.2 Azimuthal Angle Phase-Mismatched <i>S</i> -Polarization SHG Without Dispersion .....	192
3.7.2 Non-Collinear Geometry TWM With Dispersion.....	192
3.7.3 Non-Collinear Geometry FWM With Dispersion.....	196
3.8 Conclusions.....	197
4. CONCLUSIONS .....	201
4.1 Planar Waveguide Geometry Interface-Specific Optical Mixing (WISOM).....	201
4.2 Reflection Geometry Interface-Specific Optical Mixing (RISOM).....	207
BIBLIOGRAPHY.....	214
APPENDICES .....	221
CD-ROM ENCLOSURE .....	276

## LIST OF FIGURES

<u>Figure</u>	<u>Page</u>
1.1 A typical TWM experiment using two incident beams. ....	7
1.2 A typical FWM experiment utilizing three incident beams.....	8
1.3 Two $k$ -space diagrams showing (a) collinear phase-matched THG and (b) non-collinear phase-matched THG. ....	10
1.4 An adsorbate ‘bound’ to a crystal surface.....	15
1.5 A three-beam WSCARS coupling scheme. ....	17
1.6 A four coupler scheme using integrated grating couplers for four-beam mode combinations.....	17
1.7 Diagram of a typical TWM experiment in reflection geometry.....	19
2.1 A typical FWM experiment. ....	26
2.2 Structure of a planar waveguide.....	29
2.3 A cover layer of adsorbates ‘bound’ to the crystalline thin film of a planar waveguide. ....	29
2.4 Diagram of FWM in a waveguide.....	31
2.5 The electric field amplitudes of three guided waves versus depth in an asymmetric waveguide. ....	32
2.6 The product of electric field amplitudes of three guided waves versus depth in an asymmetric waveguide.....	33
2.7 Focused nearly collinear Gaussian beams crossing in a nonlinear medium resulting in a large interaction region.....	35
2.8 Focused Gaussian beams crossing at large angles yielding a small interaction region. ....	35

## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
2.9 Two possible planar waveguide coupling schemes.....	36
2.10 Diagram showing the direction angles outside the waveguide structure. ....	40
2.11 Plots of the total field versus frequency for resonances A-F and the total intensity versus surface field.....	62
2.12 Plots of the total field versus frequency for resonances A-F and the DR versus surface field.....	62
2.13 Plots of the total field versus frequency for resonances A-F and the CR versus surface field.....	63
2.14 An annotated example plot of the direct ratio.....	65
2.15 An annotated example plot of the cross ratio.....	66
2.16 The DR and CR for mode combination [2,2,0,1] at $\Delta = 1000 \text{ cm}^{-1}$ . ....	72
2.17 The DR and CR for mode combination [2,2,0,1] at $\Delta = 1400 \text{ cm}^{-1}$ . ....	72
2.18 The DR and CR for mode combination [2,2,0,1] at $\Delta = 1800 \text{ cm}^{-1}$ . ....	72
2.19 The DR and CR for mode combination [2,2,0,1] at $\Delta = 2200 \text{ cm}^{-1}$ . ....	73
2.20 The DR and CR for mode combination [2,2,0,1] at $\Delta = 2600 \text{ cm}^{-1}$ . ....	73
2.21 The DR and CR for mode combination [2,2,0,1] at $\Delta = 3000 \text{ cm}^{-1}$ . ....	73
2.22 The DR and CR for mode combination [1,0,1,1] at $\Delta = 800 \text{ cm}^{-1}$ . ....	74
2.23 The DR and CR for mode combination [1,0,1,1] at $\Delta = 1000 \text{ cm}^{-1}$ . ....	74
2.24 The DR and CR for mode combination [1,0,1,1] at $\Delta = 1200 \text{ cm}^{-1}$ . ....	74

## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
2.25 The DR and CR for mode combination [1,0,1,1] at $\Delta = 1400 \text{ cm}^{-1}$ . ....	75
2.26 The DR and CR for mode combination [1,0,1,1] at $\Delta = 1600 \text{ cm}^{-1}$ . ....	75
2.27 The DR and CR for mode combination [1,0,1,1] at $\Delta = 1800 \text{ cm}^{-1}$ . ....	75
2.28 The DR and CR for mode combination [1,1,1,2] at $\Delta = 1000 \text{ cm}^{-1}$ . ....	76
2.29 The DR and CR for mode combination [1,1,1,2] at $\Delta = 1400 \text{ cm}^{-1}$ . ....	76
2.30 The DR and CR for mode combination [1,1,1,2] at $\Delta = 1800 \text{ cm}^{-1}$ . ....	76
2.31 The DR and CR for mode combination [1,1,1,2] at $\Delta = 2200 \text{ cm}^{-1}$ . ....	77
2.32 The DR and CR for mode combination [1,1,1,2] at $\Delta = 2600 \text{ cm}^{-1}$ . ....	77
2.33 The DR and CR for mode combination [1,1,1,2] at $\Delta = 3000 \text{ cm}^{-1}$ . ....	77
2.34 The DR and CR for mode combination [0,0,0,1] at $\Delta = 800 \text{ cm}^{-1}$ . ....	78
2.35 The DR and CR for mode combination [0,0,0,1] at $\Delta = 1000 \text{ cm}^{-1}$ . ....	78
2.36 The DR and CR for mode combination [0,0,0,1] at $\Delta = 1200 \text{ cm}^{-1}$ . ....	78
2.37 The DR and CR for mode combination [0,0,0,1] at $\Delta = 1400 \text{ cm}^{-1}$ . ....	79
2.38 The DR and CR for mode combination [0,0,0,1] at $\Delta = 1600 \text{ cm}^{-1}$ . ....	79
2.39 The DR and CR for mode combination [0,0,0,1] at $\Delta = 1800 \text{ cm}^{-1}$ . ....	79
2.40 The DR and CR for mode combination [2,0,2,1] at $\Delta = 800 \text{ cm}^{-1}$ . ....	80
2.41 The DR and CR for mode combination [2,0,2,1] at $\Delta = 1000 \text{ cm}^{-1}$ . ....	80
2.42 The DR and CR for mode combination [2,0,2,1] at $\Delta = 1200 \text{ cm}^{-1}$ . ....	80



## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
2.43 The DR and CR for mode combination [2,0,2,1] at $\Delta = 1400 \text{ cm}^{-1}$ . ....	81
2.44 The DR and CR for mode combination [2,0,2,1] at $\Delta = 1600 \text{ cm}^{-1}$ . ....	81
2.45 The DR and CR for mode combination [2,0,2,1] at $\Delta = 1800 \text{ cm}^{-1}$ . ....	81
2.46 The DR and CR for mode combination [1,1,0,1] at $\Delta = 800 \text{ cm}^{-1}$ . ....	82
2.47 The DR and CR for mode combination [1,1,0,1] at $\Delta = 1000 \text{ cm}^{-1}$ . ....	82
2.48 The DR and CR for mode combination [1,1,0,1] at $\Delta = 1200 \text{ cm}^{-1}$ . ....	82
2.49 A plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at $\Delta = 1200 \text{ cm}^{-1}$ .....	91
2.50 An expanded plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at $\Delta =$ $1200 \text{ cm}^{-1}$ .....	91
2.51 A plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at $\Delta = 2000 \text{ cm}^{-1}$ .....	93
2.52 An averaged plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at $\Delta =$ $=2000 \text{ cm}^{-1}$ .....	93
2.53 A WSCARS coupling scheme for four-beam mode combinations. ....	94
2.54 Spaulding and Morris's prism-grating coupler. ....	96
2.55 Angular sensitivity to tuning for Spaulding and Morris's prism-grating coupler. ....	97
2.56 Coupling efficiency of Spaulding and Morris's prism-grating coupler. ....	97

## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
2.57 A four coupler scheme for planar waveguides using integrated grating couplers.....	98
2.58 A four coupler scheme using induced grating couplers in embedded nonlinear regions.....	101
2.59 A ring grating waveguide coupler. ....	102
2.60 Creating a ring grating exposure pattern using a circular aperture.....	103
2.61 Intensity profile of a diffractive optic for creating a ring grating exposure pattern. ....	106
3.1 Diagram of a typical TWM experiment in reflection geometry. ....	110
3.2 A semi-infinite medium comprised of layers of dipoles. ....	111
3.3 Optical mixing in a thin slab of nonlinearly active material. ....	113
3.4 A structure that gives reflected optical mixing signals for surface and bulk regions.....	114
3.5 A plot of the real part of the microscopic field versus distance into a layered dipole sheet structure. ....	121
3.6 The reflectivity of a dielectric as calculated using a layered dipole sheet structure. ....	121
3.7 A nonlinear polarization sheet at $z = 0^+$ using in the Mizrahi-Sipe Model for $P$ -polarization.....	125
3.8 Interface between the vacuum and a nonlinear medium.....	131
3.9 Generated waves at the boundary of a nonlinear medium, polarized with the electric field vector normal to the plane of reflection. ....	134

## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
3.10 Generated waves at the boundary of a nonlinear medium, polarized with the electric field vector in the plane of reflection. ....	134
3.11 Bloembergen's layered structure with its coordinate system. ....	137
3.12 Defining the wavevectors for the Bloembergen model. ....	138
3.13 A four layer structure for developing a RISOM model. ....	143
3.14 Defining the interface regions and the wavevectors of the fields. ....	146
3.15 A plot of the magnitude of the surface contribution to generated electric field for <i>S</i> -polarization collinear optical mixing in dispersionless media.....	170
3.16 A plot of the magnitude of the bulk contribution to generated electric field for <i>S</i> -polarization collinear optical mixing in dispersionless media.....	171
3.17 A plot of the DR versus incident angle for <i>S</i> -polarization collinear optical mixing in dispersionless media.....	171
3.18 An <i>S</i> -wave optical mixing RISOM experiment performed at near-grazing angles. ....	173
3.19 The <i>P</i> -wave surface amplitude versus incident angle for collinear <i>n</i> HG in dispersionless media ( $\beta = 105$ degrees). ....	175
3.20 The <i>P</i> -wave bulk amplitude versus incident angle for collinear <i>n</i> HG in dispersionless media ( $\beta = 105$ degrees). ....	176
3.21 The <i>P</i> -wave surface amplitude versus incident angle for collinear <i>n</i> HG in dispersionless media ( $\beta = 150$ degrees). ....	176
3.22 The <i>P</i> -wave bulk amplitude versus incident angle for collinear <i>n</i> HG in dispersionless media ( $\beta = 150$ degrees). ....	177

## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
3.23 The <i>P</i> -wave surface amplitude versus incident angle for collinear <i>n</i> HG in dispersionless media ( $\beta = 170$ degrees). .....	177
3.24 The <i>P</i> -wave bulk amplitude versus incident angle for collinear <i>n</i> HG in dispersionless media ( $\beta = 170$ degrees). .....	178
3.25 A plot of the DR versus incident angle for <i>P</i> -polarization collinear optical mixing in dispersionless media ( $\beta = 170$ degrees). .....	178
3.26 The <i>P</i> -wave surface contribution for collinear SHG ( $\beta = 150$ degrees), with 0.001 dispersion.....	180
3.27 The <i>P</i> -wave bulk contribution for collinear SHG ( $\beta = 150$ degrees), with 0.002 dispersion.....	180
3.28 The <i>P</i> -wave DR for collinear SHG ( $\beta = 150$ degrees), with dispersion. ....	181
3.29 The <i>P</i> -wave CR versus incident angle for collinear SHG ( $\beta = 150$ degrees), with dispersion.....	181
3.30 A non-collinear SHG experiment utilizing only the polar angles to introduce phase-mismatch. ....	183
3.31 A non-collinear SHG experiment utilizing only the azimuthal angles to introduce phase-mismatch. ....	184
3.32 Plot of the surface contribution of the field versus $\delta\theta_{1,2}$ . The process is non-collinear SHG with $\theta_{3,r,l} = 15$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	187
3.33 Plot of the surface contribution of the field versus $\delta\theta_{1,2}$ . The process is non-collinear SHG with $\theta_{3,r,l} = 45$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	187
3.34 Plot of the surface contribution of the field versus $\delta\theta_{1,2}$ . The process is non-collinear SHG with $\theta_{3,r,l} = 60$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	188

## LIST OF FIGURES (Continued)

<u>Figure</u>	<u>Page</u>
3.35 Plot of the surface contribution of the field versus $\delta\theta_{1,2}$ ( <i>sinc</i> modulation suppressed). The process is non-collinear SHG with $\theta_{3,rI} = 89$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	188
3.36 Plot of the bulk contribution of the field versus $\delta\theta_{1,2}$ . The process is non-collinear SHG with $\theta_{3,rI} = 15$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	189
3.37 Plot of the bulk contribution of the field versus $\delta\theta_{1,2}$ . The process is non-collinear SHG with $\theta_{3,rI} = 45$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	189
3.38 Plot of the bulk contribution of the field versus $\delta\theta_{1,2}$ . The process is non-collinear SHG with $\theta_{3,rI} = 60$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	190
3.39 Plot of the bulk contribution of the field versus $\delta\theta_{1,2}$ ( <i>sinc</i> modulation suppressed). The process is non-collinear SHG with $\theta_{3,rI} = 89$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	190
3.40 A plot of the ratio of total intensity ( $\times 6000$ ) versus $\delta\theta_{1,2}$ ( <i>sinc</i> modulation suppressed). The process is non-collinear SHG with $\theta_{3,rI} = 89$ degrees, $\delta\phi_{1,2} = 0$ , no dispersion. ....	191
3.41 A plot of the surface field amplitude versus $\phi_{I,I}$ . The process is non-collinear SHG with $\theta_{3,rI} = \theta_{3,\mu3} = 60$ degrees, no dispersion. ....	193
3.42 Plot of the bulk field amplitude versus $\phi_{I,I}$ . The process is non-collinear SHG with $\theta_{3,rI} = \theta_{3,\mu3} = 60$ degrees, no dispersion. ....	193
3.43 A plot of the polar angular intensity profile versus ‘reflection’ angle. The process is Gaussian beam SFG with dispersion. ....	195
3.44 A contour plot of the intensity angular profile versus the polar and azimuthal angles. The process is Gaussian beam CARS-like FWM with dispersion. ....	197

## LIST OF APPENDICES

<u>Figure</u>	<u>Page</u>
APPENDIX A .....	222
A.1 The Wave Equation and Boundary Conditions for Inhomogeneous Waveguides.....	222
A.2 Field Solutions for Asymmetric Step Index Dielectric Waveguides.....	235
A.3 The Intensity, Phase-Mismatch, and Quantitative Tools for WISOM.....	244
APPENDIX B.....	249
B.1 Determining the Solutions to the Inhomogeneous Wave Equation.....	249
B.2 Explicitly Including the Phase-Mismatch Information Between Layers .....	257
B.3 Determining the Electric Fields in the Layered Structure .....	260
B.3.1 Determining the <i>S</i> -polarized Electric Fields in the Layered Structure .....	262
B.4 Determining the <i>P</i> -polarized Electric Fields in the Layered Structure .....	269
B.5 Connections to Bloembergen's Results .....	273
B.6 Summary of the Field Expressions for the Macroscopic RISOM Model .....	274

## LIST OF APPENDIX FIGURES

<u>Figure</u>	<u>Page</u>
A.1 A cylindrical Gaussian volume straddling an interface.....	225
A.2 A Stokesian loop straddling an interface.....	227
A.3 One of the beam coordinate systems and the (unprimed) waveguide coordinate system. ....	235
B.1 A polarization and propagating field at $(\tilde{r}', t')$ and the displacement field observed at $(\tilde{r}, t)$ .....	251
B.2 Angle definitions for the nonlinear polarization, the source and generated wavevectors, and generated electric field.....	255
B.3 Diagrammatic aid in determining the phase-mismatch in a nonlinear layer. ....	259

## DEDICATION

To those whom stone houses build:

שתי אבנים בונות שני בתים  
שלש אבנים בונות ששה בתים  
ארבע אבנים בונות מאה ועשרים בתים  
שש אבנים בונות שבע מאות ועשרים בתים  
שבע אבנים בונות חמשת אלפים וארבעים בתים  
מכאן ואילך צא וחשוב מה שאין הפה יכול לדבר ואין האוזן  
יכולה לשמוע:

שפר יצרח א:טו

*Two stones build 2 houses,  
Three stones build 6 houses,  
Four stones build 24 houses,  
Five stones build 120 houses,  
Six stones build 620 houses,  
Seven stones build 5040 houses.  
From here on go out and calculate that which the mouth cannot speak  
and the ear cannot hear.*

*(Sefer Yetzirah 1:16)*



# THEORETICAL ANALYSIS OF OPTICAL MIXING AT INTERFACES

## 1. INTRODUCTION

Can vibrational and electronic surface resonances involving optical transitions be detected on any flat surface? This is the central question of this research. This research was driven by the potential to further develop the tools of surface science. A goal of surface science is to understand the physics and chemistry of the region only a few atomic layers thick between a vacuum and deep within the structure of a solid. Often the physics of interest involves characterizing the electronic or vibrational resonances of the constituent species of that thin interstitial region called the 'surface'. To accomplish this requires probes sensitive to this region. Separating the resonances of the 'surface' from the resonance behaviour of the rest of the solid (i.e., the bulk) is rarely trivial. It is the aim of this research to add to the continuing development of probing techniques that involve *interface-specific* optical mixing (ISOM).

Why should optical mixing (OM) be considered over techniques commonly utilized by commercially available surface science units? There are many well-established techniques that can be used to probe a surface. With optical mixing being the exception, virtually all of these techniques involve electron-electron scattering (e.g., LEEDS or AES), electron-photon scattering, or photon-electron scattering (e.g., XPS or UPS). The apparatus are highly specialized for each technique and require the subject surface be under ultra-high vacuum conditions. The greater flexibility of optical mixing in the visible region allows the probing of more states of a system with a single apparatus. It also allows for the simultaneous probing of the energy and symmetry of resonant transitions. The long mean free path of light allows the technique to be used from vacuum to atmospheric conditions. By virtue of long path length, ISOM can even be done at an interface between two media deep inside a (transparent) sample. Unfortunately, because of the long penetration depth,

there is no apparent preference to interface-generated over bulk-generated optical signals. Thus, nonlinear optical processes are not inherently interface-specific. However, by choosing the geometries of the applied fields, the generated fields can have enhanced contributions from the surface over that of the bulk. Commonly used optical mixing techniques such as second harmonic generation (SHG) and sum parametric mixing (SPM) utilize geometries that embellish the ratio of the surface to the bulk contribution. Much of the work presented here involves exploring various other ways to manipulate the applied fields in order to obtain favourable ISOM conditions. Here, two specific methods of achieving ISOM are developed.

Chapter 2 takes the approach of manipulating the applied fields to achieve an interference condition everywhere except the surface region; hence optical mixing occurs only in the surface region. This is done in the confines of a planar waveguide. This document refers to ISOM in a waveguide geometry as waveguide ISOM (WISOM). The research performed here extends the information available in the literature by creating a data base of useful WISOM parameters and conditions. This is done for specific OM processes in a waveguide. Phase-matching is considered in the waveguide geometry as a way to increase the surface signal. Phase-matching inside a waveguide has never been reported in the contemporary literature. Of the several conclusions drawn, one is that to experimentally scan through the optical surface resonances several waveguides would be required. This would mean changing waveguides during the experiment, which is undesirable. Another conclusion is that use of phase-matching in a waveguide increases the overall signal level and yields a better signal-to-noise ratio. Increasing the signal levels via phase-matching, however, does not come without expense. The experimental difficulty is considerably increased. Due to the difficulty of WISOM experiments, a second approach was considered.

In Chapter 3, reflection geometry ISOM (RISOM) is considered. Manipulation of the applied fields to maximize the surface-to-bulk signal ratio is done by selecting the angles

and polarizations of the incident waves. The angular dependence of optical mixing and phase-matching conditions are analyzed as a way to predict enhanced surface-specificity. Advantages of the RISOM technique include a straightforward experimental geometry, typically large signals, and no apparent limitations on the type of subject surface. Results show that there exists a nonlinear Brewster's angle analogous to the familiar linear Brewster's angle. A wave reflecting off a surface near Brewster's angle has no  $P$ -polarization component. This condition holds for both the linear and nonlinear Brewster's angles. It is shown that for dispersive media, the nonlinear Brewster's angles are different for surface and bulk. It is predicted that the use of this angular separation provides one enhancement of the surface-to-bulk signal ratio. Furthermore, enhancements can be made by utilizing the different angular dependencies of optical mixing in the surface and bulk signals. Chapter 3 is considered the most unique and important part of this research. It comprises the larger portion of this document. In order to predict RISOM experiments, an entire treatment of optical mixing at a surface that includes both surface and bulk contributions is developed. The model developed in this chapter is an exact macroscopic treatment of optical mixing at an interface and is hitherto undocumented in the literature. An extensive analysis is done in an attempt to understand the angular and polarization dependencies of the bulk and surface generated waves. This complete macroscopic formulation should be considered a *starting point* for both new experiments and future theoretical research.

This introductory chapter provides perspective into optical mixing in general and ISOM specifically. First an overview of this research is outlined. This research developed as it progressed; as each question was investigated, new questions arose which changed the path of the programme. After the overview outlining the directions this research eventually took, a brief history and review of established surface science probes is presented. With these perspectives, a more detailed outline of each aspect of this research is presented.

## 1.1 How This Research Developed

The original goal of this research was to experimentally probe the electronic states of simple molecules on a surface. The question was: what can be learned about the mechanisms that cause a simple (two atom) molecule to ‘stick’ to a surface by studying how the electronic states of the molecule change? It was proposed that this research be coupled to the research of others who might characterize the vibrational states of such molecules. Four-wave ISOM performed in dielectric waveguides (four-wave WISOM) would provide the surface-specificity needed to observe less than a monolayer of the subject molecules on a surface. This technique has been used to characterize vibrational states of  $\text{O}_2^-$  [1], ethylene [2], phenol and pyridine [3], and many other species on zinc oxide waveguide surfaces. After an initial theoretical analysis of this technique for *electronic* states, the experiment appeared to have little promise due to low signal levels, hard-to-build and cumbersome-to-use waveguides, and the necessity of having to change waveguides during the course of the experiment. At the time, the prevailing literature mainly considered combinations of waveguide modes using three beams to achieve the interference condition necessary for surface-specificity. Furthermore, no researcher had reported phase-matching in the waveguide as a way to increase signal levels. An extensive theoretical investigation ascertained if using four-beam waveguide modes and possibly phase-matching would make WISOM a viable technique for the study of electronic resonances. At the conclusion of this investigation it was realized that there are four-beam waveguide modes that provide broader tuning ranges while allowing for phase-matching. Unfortunately, these four-beam waveguide modes increase the experimental difficulty. For instance, performing surface CARS in a waveguide (WSCARS), requires two couplers and three beams, one beam carrying two waves with degenerate wavevectors. Using four-beam modes, the degeneracy is broken and four couplers are needed to couple four different waves to the waveguide. It is experimentally troublesome to separate one

beam to create two non-degenerate waves, to get the timing of these beams correct, and to have them overlap in the waveguide. Moreover, it is very difficult to construct four different couplers on a single waveguide. Using prism couplers, it is hard enough to get two couplers on a waveguide, let alone four. Attempts were made to integrate grating couplers into waveguides with limited success. Again the problem of constructing four such integrated couplers proved too laborious. Suggestions on the use of diffractive optics to construct ring grating couplers were made. Experimentation showed this approach may be valuable, but would require much more research to be realizable.

The theoretical analysis of utilizing waveguides for electronic and vibrational resonance ISOM provided one lamentable conclusion: WISOM is experimentally challenging. A new question was considered: is there a more experimentally straightforward technique? The idea of using a simple reflection geometry ISOM (RISOM) was considered. However, an extensive literature search revealed no RISOM documented techniques utilizing *general* three- or four-wave optical mixing. There is a wealth of literature on use of second harmonic generation (SHG) on the surfaces of materials to achieve ISOM. Realizing that, just as in the linear case, the reflection of a nonlinear signal from an interface has angular dependencies, a new question was posed: would it be possible to utilize these dependencies to achieve RISOM? To answer this question, a model was needed that could predict the relative strengths of the signals generated from surface and bulk. Without a model in the literature to predict RISOM, a new model was constructed. This model is an extensive modification of an existing model originally developed by Bloembergen to describe optical mixing at dielectric surface. The Bloembergen model, however, does not include contributions from the bulk of the dielectric. With this new model, optical mixing from *both* the surface and bulk regions of a medium could be compared. This exact macroscopic model, which correctly handles possible deviations from perfect phase-matching, was used to suggest combinations of reflection geometries that would lead to RISOM experiments. After completing these two

theoretical analyses, it was apparent that the scope of the work shifted from experimentally investigating out electronic states of simple molecules on a surface to making fundamental contributions to the use of ISOM as a surface science probe.

## 1.2 Review of Optical Mixing and Definition of Optical Mixing Terms

Optical mixing has been used as a form of spectroscopy since shortly after the advent of the laser in the early sixties. Soon after frequency doubling, tripling, and optical frequency mixing processes were first observed ([4], [5], [6], [7], [8]), the use of these processes for obtaining interface-specific information has developed. For instance, second harmonic generation (SHG) on crystals with inversion symmetry necessarily leads to interface-specific signals under the dipole approximation. This is by virtue of the breaking of that symmetry at an interface. This powerful ISOM technique has been reviewed in a recent paper by Corn and Higgins [9]. Despite the successes of surface-specific SHG, there is a limit to its application: only centro-symmetric surfaces or surfaces of materials with insignificant quadrupole moments can be studied. Much of the motivation of this research was to investigate ways which would allow the study of a less limited class of flat surfaces. Moreover, additional information about optical surface resonances can be found using general three-wave mixing (TWM), four-wave mixing (FWM), or higher order processes --  $n$ -wave mixing ( $n$ WM). Note that SHG is a second order process, which is a special case of TWM. These other processes are important in the search for ISOM experiments that will yield more information about surface resonances than SHG alone.

As can be seen in the previous paragraph, there are a number of acronyms common in nonlinear optics. It is warranted to remind the reader before proceeding too deeply into the morass of acronyms the most common of these. In this document, there are general and specific nonlinear processes. For instance, three-wave mixing (TWM) and four-wave mixing (FWM) are general OM processes commonly referred to. These general processes governed by the second and third order nonlinear susceptibilities, respectively. Thus,

there must be three waves at  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$  involved in TWM. In FWM, there are four waves at  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$ , and  $\omega_4$ . Even  $n$ -wave mixing ( $n$ WM) can be defined as that process that is governed by the  $(n-1)$ -th order nonlinear susceptibility. Note that no relationship has been imposed on the incident beams carrying these waves. There may be a single incident beam carrying all  $n-1$  waves, there may be several beams, or there may be  $n-1$  incident beams each carrying one of the waves that are mixed to give the  $n$ -th wave in a given  $n$ WM process.

Specific cases to TWM are SHG, sum frequency generation (SFG), and difference frequency generation (DFG). An example of a TWM experiment is given in Figure 1.1. Here SFG is shown in medium with a linear (first order) and a second order susceptibility. Two waves enter the material and a third wave is generated. The generated wave has frequency  $\omega_3$  and is related to the incident waves at  $\omega_1$  and  $\omega_2$  by  $\omega_3 = \omega_1 + \omega_2$ , where  $\omega_1 \neq \omega_2$ . If  $\omega_1 = \omega_2$ , the process is SHG (the two incoming waves have degenerate wavevectors and are contained in a single beam). When the generated wave has frequency  $\omega_3 = \omega_1 - \omega_2$  or  $\omega_3 = \omega_2 - \omega_1$ , the process is known as DFG.

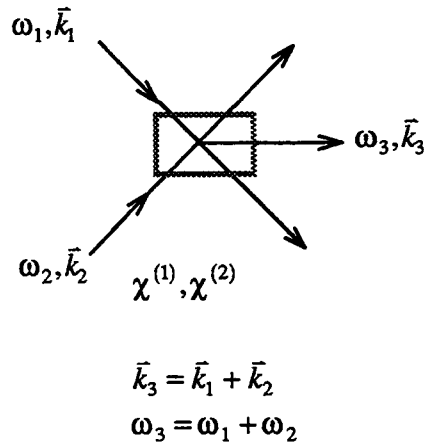


Figure 1.1: A typical TWM experiment using two incident beams.

There are many possible special cases of FWM; many do not have names. There is only one acronym used in this document that involves a specific case of FWM. This is CARS (coherent anti-Stokes Raman spectroscopy). Since CARS refers to vibrational resonances, and it is possible that electronic resonances are meant, the term CARS-like FWM is used. Throughout this document CARS and CARS-like FWM are used interchangeably; CARS-like FWM is always understood. CARS-like FWM experiments are done in a material having a non-zero third order susceptibility and are typified by  $\omega_4 = \omega_1 - \omega_2 + \omega_3$ , where  $\omega_1 = \omega_3$ . The geometry of a such an experiment is shown in Figure 1.2.

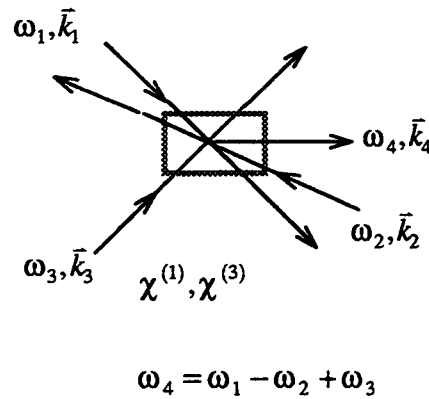


Figure 1.2: A typical FWM experiment utilizing three incident beams.

For any OM process, the difference in phase (the phase-mismatch) between the induced source polarization and the generated traveling wave must be considered. This is a particularly important concept as will be seen later. Thus, it is reviewed at this juncture. The wave generated in an OM process can be thought of as originating from an oscillating source polarization induced by the action of the incident waves on the nonlinear medium. The induced polarization can be considered the tensor product of the nonlinear (electric) susceptibility and the incident (electric) fields. For example, for FWM, the (third order)



polarization is given by

$$\bar{P}^{(3)}(\omega_4) = \chi^{(3)}(\omega_4; \omega_1, \omega_2, \omega_3) \bar{E}(\omega_1) \bar{E}(\omega_2) \bar{E}(\omega_3) e^{i[(\bar{k}_1 + \bar{k}_2 + \bar{k}_3) \cdot \bar{r} - \omega_4 t]}.$$

It is recognized that the source polarization has a phase  $(\bar{k}_1 + \bar{k}_2 + \bar{k}_3) \cdot \bar{r}$ . Note that it is assumed that the medium through which each wave travels has dispersion. The indices of refraction appear explicitly in the source polarization phase. For instance for wave 1, the index of refraction is  $n_1 = n(\omega_1)$ . Thus, the induced nonlinear polarization is more precisely written as

$$\bar{P}^{(3)}(\omega_4) = \chi^{(3)}(\omega_4; \omega_1, \omega_2, \omega_3) \bar{E}(\omega_1) \bar{E}(\omega_2) \bar{E}(\omega_3) e^{i[(\bar{k}_1(\omega_1) + \bar{k}_2(\omega_2) + \bar{k}_3(\omega_3)) \cdot \bar{r} - \omega_4 t]}.$$

For FWM,  $\omega_4 = \omega_1 + \omega_2 + \omega_3$ , but if the medium has dispersion, in general,  $\bar{k}_1 + \bar{k}_2 + \bar{k}_3 \neq \bar{k}_4$ . This is true of mixing processes in single-beam, that is collinear wave, geometries. The generated wave can be thought of as having a phase,  $\bar{k}_4(\omega_4) \cdot \bar{r}$ . The field of the generated wave can be represented by:

$$\bar{E}(\bar{r}, t) = \bar{E}(\omega_4) e^{i[\bar{k}_4(\omega_4) \cdot \bar{r} - \omega_4 t]}.$$

Using the oscillating polarization as a source term in Maxwell's equations yields a generated wave that depends on the phase difference:

$$\bar{E}(\omega_4) \sim e^{i[(\bar{k}_1(\omega_1) + \bar{k}_2(\omega_2) + \bar{k}_3(\omega_3) - \bar{k}_4(\omega_4)) \cdot \bar{r}]}.$$

The phase of the source polarization is generally not equal to the phase of the generated wave due to dispersion. Only if the directions of the wavevectors  $k_1$  through  $k_3$

are properly adjusted can there be no phase difference, that is no *phase-mismatch*. Zero phase-mismatch occurs when the generated wave oscillates with the same phase as the polarization induced by the incident waves. The reader is reminded that when there is no phase-mismatch (when phase-matched), the generated signal becomes very large.

There are two particular cases of phase-matching: collinear and non-collinear. If there is no dispersion, phase-matching is collinear. That is, phase-matching is done with all waves collinear. In this situation, it is seen that  $|\vec{k}_1| + |\vec{k}_2| + |\vec{k}_3| = |\vec{k}_4|$  since all the indices of refraction are equal. The phases of the polarization and generated wave are equal when all the waves travel in the same direction. This is illustrated in the  $k$ -space diagram in Figure 1.3 (a). In the case of a non-zero dispersion, non-collinear phase-matching is used. The finite dispersion requires that phase-matching be done by adjusting the directions of the incident waves as is indicated in Figure 1.3 (b).

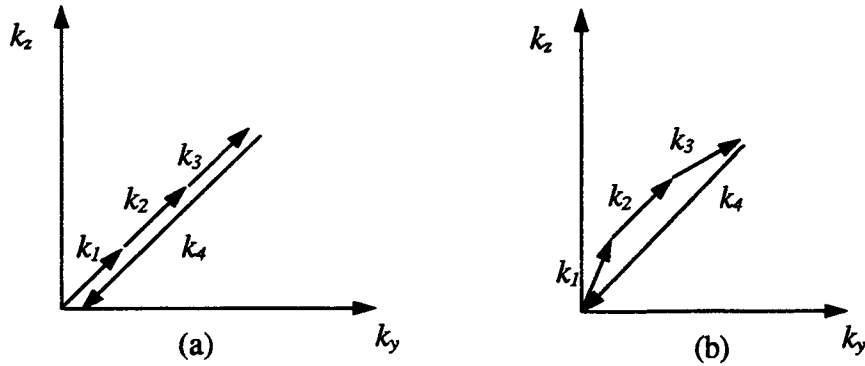


Figure 1.3: Two  $k$ -space diagrams showing (a) collinear phase-matched THG and (b) non-collinear phase-matched THG.

It is always possible to introduce an intentional phase-mismatch in an otherwise phase-matched OM process. This is done by adjusting the direction of the wavevectors away from the phase-matched directions. As will be seen throughout this document,

introducing an intentional phase-mismatch can be utilized to reduce the OM contribution from an unwanted region of a ‘surface’. This allows for enhanced surface-to-bulk signal ratios when performing ISOM experiments.

### 1.3 Interface-Specific Optical Mixing (ISOM)

As the electronics and related industries move to utilization of nano-structure devices, surface science is finding an ever increasing role in the commonplace. A goal of surface science is to develop a fundamental understanding of the physics and chemistry at specific interfaces. This is done by probing the surface states at the interface. Characterizing a surface may mean mapping out the states of the surface in order to describe the surface physics. Using surface probes, the surface scientist may track changes in surface states as external changes are made to the surface. Surface chemistry may be studied as well as surface dynamics with surface-specific probes.

What is a surface and how are they studied? Surface scientists are armed with a battery of techniques to probe surfaces, but to understand these techniques, a portrait of a surface must be painted. A surface may be described in many ways depending on one’s perspective. Even if limited to the “perfect crystal surface”, there are a myriad of ways to view the interface between a vacuum (or atmosphere) and a bulk crystal. We tend to have a naive idealization of an interface; the experimenter’s “perfect crystal surface” actually includes an interstitial region comprised several of atomic layers. Within this interstitial region there may be a relaxation of the unit cell parameters. As a result, the response to external electric and magnetic fields (probes) is different than that of the bulk crystal. Physical defects (e.g. dislocations and steps) occur in this region; the “perfect crystal surface” can only be spoken of in a theoretical sense. Impurity concentrations may increase as one moves further from the native crystal to the other side of the interface. Such impurities may be caged in the lattice, while others may be substituted for native atoms in the lattice. Above this region, having been scavenged from the atmosphere by

the crystal's dangling bonds, molecular and atomic species reside; these too effect the response to externally applied electric and magnetic fields used as surface probes.

This surface region is to be contrasted with the rest of the solid. The structure deep inside a solid crystal is considered periodic on a large scale (thousands of lattice constants in dimension). Generally, the physical interaction of an applied field within a crystal is determined by the identity, concentration, and geometrical arrangement (symmetry) of the chemical species comprising the matter. When a surface is present, there are additional interactions due to the rather different morphology of the surface versus the bulk. The aim of surface-specific experiments is to separate the interactions due to the presence of a very large number of atoms comprising the bulk crystal from the surface interactions. Consider that the surface density of surface atoms is typically  $10^{15}/\text{cm}^2$ . From Avogadro's number, the volume density of the bulk atoms is  $10^{23}/\text{cm}^3$ . In a typical experiment, the probed volume is  $10^{-8}\text{cm}^3$  accomplished by means of a beam focused to a  $100 \times 100$  micron spot penetrating 1 micron (about one wavelength of light) deep. In this case, the number of surface atoms is of the order of  $10^{11}$ . This is to be compared to  $VN_A = 10^{15}$  atoms sampled in the bulk. Thus, a small surface-derived signal rides atop an enormous bulk background signal. Hence, great effort must be placed into surface probing techniques in order to extract the surface information.

Many surface science probes detect electrons ejected or scattered from surfaces. By virtue of the short mean free path of electrons through a solid, electrons ejected from a solid almost certainly originate from the surface region. In scattering experiments, the small penetration depth of electrons into the surface guarantees that the states they probe be surface-related. In these ways, techniques such as x-ray photoelectron spectroscopy (XPS), ultraviolet photoelectron spectroscopy (UPS), Auger electron spectroscopy (AES), electron energy-loss spectroscopy (EELS), and low-energy electron diffraction spectroscopy (LEEDS) are interface-specific. Because these techniques rely on scattered particles, they are limited to high and ultra-high vacuum environments. The mean free

path of an electron in air is only millimeters. This is a great disadvantage as few real world surfaces exist in UHV conditions. In realistic environments, such as atmospheric conditions, these probes are unsuitable; *ex situ* studies must be performed. Such studies must be done with some circumspect. Between measurements of the sample, the surface experiences transfers in and out of vacuum chambers and perhaps other undesirable or unavoidable changes. Another disadvantage with these scattering techniques is that alone, they can only give limited state information. Several of these techniques may be required to map out surface states of at an interface. For instance, XPS and UPS are only sensitive to the states of the core and valence electrons, whereas LEEDS can only be used to detect the symmetry of surface states.

Why should surface scientists be interested in ISOM? ISOM offers greater flexibility as it allows the probing of more states of a system using a single apparatus. The long mean free path of light allows ISOM-based techniques to be used in virtually any environment from atmospheric to vacuum conditions. The long path length also permits ISOM to be done at an interface between two media deep inside an optically transparent sample. Unfortunately, there is no preference to interface-specific signals over bulk-generated optical signals due to the long mean free path of light.

Nonlinear optical processes can be made surface-specific by particular choices of symmetry or geometry. The classic example of ISOM is SHG. Under the dipole approximation, SHG is forbidden in a crystal with inversion symmetry. Thus, a centrosymmetric crystal will only produce a SHG signal at an interface, where this symmetry is broken. It should be noted that SHG is sensitive to both to symmetry and the frequencies of electronic resonances of a surface. (See for example: [10], [11], [12], [13], [14], [15], [16], [17], [18], [19], [20], [21].) A general solid is not centro-symmetric. Thus, there is no apparent reason for an optical mixing experiment to be interface-specific. There are two possible approaches to obtaining ISOM. One is to engineer the incident fields such that the generated field is non-zero everywhere but in the surface region. This approach is

utilized in WISOM. The other is to rely on the geometries and polarizations of the applied fields and angular dependencies of the generated fields in an attempt to enhance the surface over the bulk contribution to the total signal. This is embodied in RISOM. An intense theoretical analysis of these two approaches to ISOM is the major subject of this research.

#### 1.4 ISOM Using Planar Waveguides (WISOM)

For vibrational resonances, waveguide surface CARS (WSCARS) has been proven to yield favourable ISOM conditions. WSCARS (i.e., CARS-like WISOM) can provide the surface-specificity needed to observe less than a monolayer of the subject molecules on a surface. With the desire to study not the vibrational states of such molecules but the electronic resonances, an initial theoretical analysis of this technique was done for *electronic* states. The initial analysis showed that any WISOM experiment would be ineffective for four reasons: low signal levels, having to deal with cumbersome waveguides, including the difficulty of constructing waveguides, and necessity of having to change waveguides during the course of the experiment. Construction of a waveguide requires precise control of both the index of refraction and the thickness of the guiding thin film. Typically, many attempts of fabricating a waveguide with specific parameters must be made before a satisfactory one is produced. Thus, during a given experiment, different waveguides with varied thicknesses and film indices must be exchanged as the field frequencies are scanned. To complicate matters, a single waveguide is effective over only limited tuning ranges. In a way, Chapter 2 explores how to best choose a waveguide for a given experiment. Before this work was begun, the prevailing literature considered only three-beam combinations of waveguide modes to achieve the interference condition necessary for surface-specificity. Furthermore, there was no documentation on the use of phase-matching in a waveguide as a way to increase signal levels. Phase-matching and enhanced interference conditions made possible by *four-beam* mode combinations allow the fullest exploitation of a given waveguide for a WISOM experiment.

From this initial investigation, different questions became important. Is there a way to maximize the surface-to-bulk signal for a WSCARS experiment over what is commonly done? Can the resonances of an adsorbate at a surface be separated from the resonances of the bulk crystal? A crude illustration of this is given in Figure 1.4. An adsorbate adheres to a surface and has vibrational and electronic resonances. The crystal itself has inherent vibrational and electronic states. Because there are many fewer adsorbate species than atoms in the crystal, the task of separating out the affects of the adsorbate resonances over the large background of crystal resonances is daunting.

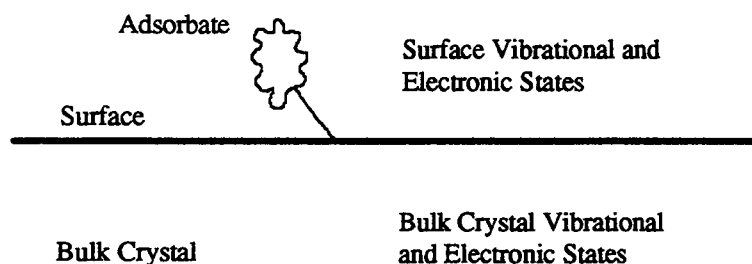


Figure 1.4: An adsorbate ‘bound’ to a crystal surface.

An extensive theoretical investigation was launched to map out all experimentally relevant parameter space in order to ascertain if using four-beam waveguide modes and possibly phase-matching would make WISOM a more accessible technique. The usual approach of solving for the generated field was adopted: First, the (range of) frequencies of the incident waves are picked. Then a mode combination is chosen; one with the potential of satisfying the interference condition leading to surface-specificity. Performing WSCARS in the three-beam way requires two couplers and three beams, one beam carrying two waves with degenerate wavevectors. This means two of the waves are

coupled into the same waveguide mode. In order to meet the interference condition, the product of the (field profiles for each of the) modes must be even. Using the three-beam mode combinations, mode combinations  $[0,0,0,1]$ ,  $[0,1,0,0]$ ,  $[1,0,1,1]$ , and  $[1,1,1,0]$  lead to an interference condition. Here the mode combinations are labeled as: [mode of wave with  $k_1$ , mode of wave with  $k_2$ , mode of wave with  $k_3$ , mode of wave with  $k_4$ ]. Note that the mode of the first and third waves are always equal for these three-beam mode combinations. Second, the guiding film thickness and index of refraction are adjusted to maximize the interference condition. Often, the interference condition is poorly met and the mode combination is discarded as non-viable. Once maximized, it is seen that the interference condition is maximum for specific waveguides with certain mode combinations. Furthermore, the range of frequencies (i.e., the tuning range) over which the interference condition is favourable is found to be rather limited for any given mode combination and waveguide. Also, the phase-mismatch of the mode combinations is large if  $k_3 = k_1$ . Hence, the signals are small. This is especially so if the film has dispersion.

This inquiry gave a new direction to the analysis: can phase-matching be done in the waveguide as a means of increasing the signal levels? In order to achieve phase-matching, the degeneracy of  $k_1$  and  $k_3$  must be broken. This is done by separating the beam carrying the waves with  $k_1$  and  $k_3$  into two beams. Hence the FWM process is done with a total of four beams. With these beams separated, there is no reason to have them in the same mode. In fact, placing them in different modes is advantageous from the standpoint of maximizing the interference condition. Hence, four-beam mode combinations are considered. To understand the experimental difference between the three-beam and four-beam mode combinations, consider Figure 1.5 and Figure 1.6. Figure 1.5 shows a WSCARS experiment using three-beam mode combinations where wave 1 and wave 3 are carried by the same beam. The experiment requires two couplers and three beams as is shown. Using four-beam modes, four couplers are needed to couple four different waves to the waveguide. Such a coupling scheme is illustrated in Figure 1.6. Here grating couplers are utilized instead of prism couplers. Clearly, experiments utilizing four-beam



mode combinations are more cumbersome. Despite the experimental complexities, it was thought that improvements in WISOM could be had by employing the four-beam combinations while phase-matching inside the waveguide.

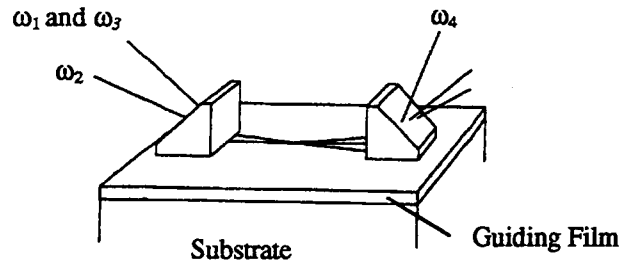


Figure 1.5: A three-beam WSCARS coupling scheme.

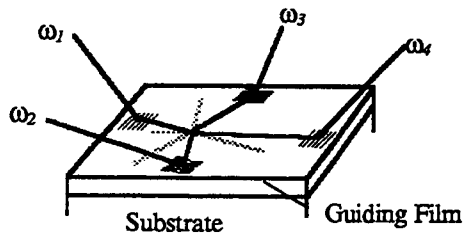


Figure 1.6: A four coupler scheme using integrated grating couplers for four-beam mode combinations.

From the theoretical analysis several conclusions are drawn. Most importantly, the use of four-beam mode combinations allows for both favourable interference conditions and phase-matching in waveguides. Improvements over the surface-to-bulk signal ratio can be as large as a factor of 100, and are typically larger than 10. The overall signal levels, because of the phase-matching, can be increased by factors ranging from 10 to 10,000. Furthermore, the tuning range (the range over which a single waveguide exhibits favourable interference conditions for a given mode combination) is increased for some

conditions. Despite this improvement, to experimentally scan through the available surface resonances, several waveguides would still be required. This would mean changing waveguides during the experiment, which is undesirable. Despite the additional difficulty, four-beam mode combination experiments seem to be advantageous. Calculations show that typically to achieve phase-matching, the angle between each beam needs to be approximately 90 degrees regardless of the mode combination. To offset this difficulty, sections are included on waveguide coupling using gratings. Fabrication techniques for integrated grating couplers are discussed. Techniques for building grating structures are suggested based on theoretical inquiry. These include constructing ring gratings by masking diffraction from circular apertures and diffractive optics generated by computer comprised by a superposition of  $J_0$  functions.

The final conclusion that is reached is that WISOM is an experimentally difficult technique. Furthermore, there are limitations on the kind of surfaces that can be studied. Only *dielectric* surfaces for which *thin films can be grown* on dielectric substrates can be scrutinized. These thin film surfaces must have relatively high indices of refraction capable of supporting guided modes. Finally, the reproducibility of growing these film surfaces should be good, since a battery of waveguides must be constructed in order to scan through the resonances of the surface.

## 1.5 ISOM Using Reflection Geometry (RISOM)

With the limited success of WISOM, the old question seemed still relevant: Can the surface-to-bulk signal ratio be maximized in ISOM experiments? Since there is a small surface signal riding atop a large bulk signal (which is considered noise in a surface resonance experiment), how can the typically poor signal-to-noise ratio be maximized? Chapter 2 takes the approach of using ISOM in waveguide geometry to maximize the surface-to-bulk signal ratio. The research reported in Chapter 2 extends the information available on WISOM, but leaves the question with a limited and unsatisfactory answer.

Attempting to ask a more specific question, a new query was posed: Can simple *reflection geometry* be employed in an optical mixing process and still be used to detect a surface component to a signal that is composed mostly of bulk contributions? A reflective geometry experiment is straightforward and phase-matching is easy to achieve; signals are typically large and there are no apparent limitations on the type of surface that can be studied. A typical experiment is shown in Figure 1.7. Specifically, it shows a three-wave mixing process in reflection geometry. Contrasted with the waveguide geometry, in this geometry, experiments are relatively easy. There is no need for coupling into a guiding structure. However, in the waveguide geometry, there is an engineered interference condition; the reflection geometry has no apparent interference condition. It is the interference condition in WISOM that prevents signals from being generated everywhere but the surface monolayer and gives superb surface-to-bulk signal ratios.

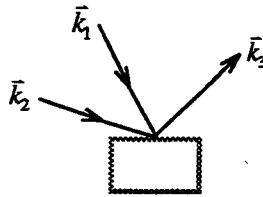


Figure 1.7: Diagram of a typical TWM experiment in reflection geometry.

What conditions yield favourable signal-to-noise ratios in RISOM experiments? In a linear process, the reflection off a semi-infinite medium obeys the (linear) law of reflection. The angular and polarization dependencies are described by the familiar (linear) Fresnel equations. By analogy, when the process is nonlinear, reflections obey *nonlinear* Fresnel equations that describe their angular and polarization dependencies. Since the surface and bulk are very different as discussed in Section 1.3, the nonlinear Fresnel equations for the surface and bulk are dissimilar. This means the angular and, perhaps, polarization behaviour is different for the two regions. On this basis, it is possible to separate the

surface and bulk signals in a nonlinear optical process that employs reflection geometry. How then, can RISOM experiments be theoretically predicted?

An exhaustive literature search revealed no theoretical models that can simultaneously predict the surface and bulk contributions of an OM process in reflective geometry. Undoubtedly the best approach to creating such a model is a microscopic dipole sheet model. Such a model has been used to calculate the microscopic local fields *inside* a semi-infinite dipole structure where all the dipoles were identical [22]. It proved very successful, but involved. Rather than do this time-consuming and arduous calculation, a less precise but much simpler macroscopic model was adopted. In a series of papers, Bloembergen and co-workers determined the theoretical equations that govern optical mixing (particularly SHG) in reflection geometry [23], [24], [25]. This *macroscopic* approach assumed optical mixing occurred within a thin slab, or surface region, of nonlinear material. Below the surface layer is a semi-infinite bulk with only linear properties. The linear properties of the bulk are chosen to be different than the surface slab. The interface between the surface and bulk provides a mechanism for the generated light to reflect and return to the vacuum. The surface-to-bulk index ratio appears in the nonlinear Fresnel equations as a scaling factor in the amplitude of the generated field as measured in the vacuum.

A new, macroscopic model was constructed to predict the first favourable RISOM conditions. An exact microscopic model was considered too fatiguing for first-time predictions. The new model was constructed by extensive modification of Bloembergen's macroscopic model. His model, however, does not include OM contributions from the bulk which must also have nonlinear properties. Beginning with Bloembergen's model, an additional layer is added *between* the surface slab and the bulk. This layer is referred to as the *bulk median*. This layer is given nonlinear properties different from that of the surface. The bulk median's linear index of refraction also differs from the semi-infinite bulk. This new interface establishes a mechanism for the light generated in the bulk (median) to be

reflected back into the surface layer and into the vacuum. This bulk-bulk median interface provides the same reflective boundary as the bulk-surface interface does in the Bloembergen model, hence the new RISOM model simultaneously predicts both nonlinear surface and bulk signals. The ratio of the indices of refraction between bulk median and bulk become scaling parameters in the nonlinear Fresnel coefficients. A similar argument is made for the thickness of the bulk median and surface layers. Since the layers are introduced arbitrarily, the thickness is an arbitrary scaling parameter that appears primarily in the phase of the field originating from the two layers. With light re-entering the vacuum generated from *both* the surface and the bulk (median), a comparison of the ratio of amplitudes can be made. The field amplitudes are derived as they are in the Bloembergen model: all the incident, reflected, and transmitted fields are defined in the various layers and the boundary conditions are satisfied at the interfaces between layers. Analytical expressions for the bulk and surface generated field amplitudes are then found. These expressions contain the incident field amplitudes, frequencies, directions, and polarizations as well as the reflected waves' directions and polarizations. The actual derivation of these expressions is detailed in Appendix B. The result is a more precise model for interface-specific optical mixing.

The model is then used to predict ISOM conditions for specific processes such as CARS-like FWM, SHG, and TWM. The major portion of Chapter 3 is devoted to mapping out the parameter space of OM in reflection geometry. Specifically, the model is used to find conditions that yield enhanced surface-to-bulk signal ratios, that is, favourable RISOM conditions. First only collinear phase-matched geometries are considered. Collinear phase-matching is appropriate for dispersionless materials as well as harmonic generation. Both polarization states,  $S$  and  $P$ , are explored. In general, it is found that improvement in the surface-to-bulk signal ratio is possible for  $S$ -polarization when the reflection angle is large; for  $P$ -waves, enhancement is done by making use of the nonlinear Brewster's angle -- the nonlinear analogue to the linear Brewster's angle. Non-collinear phase-matching is then considered; it is appropriate when the linear dispersion of materials

is important. The angular dependencies for the two polarization states for SHG (where dispersion is very important, but phase-matching is automatic) and CARS-like FWM (where dispersion is less important, but phase-matching is crucial) are specifically considered. It is found that the conditions for enhanced signal-to-noise ratio for collinear and non-collinear phase-matching are identical. Enhancement is generally better, however, when the surface and bulk dispersions are dissimilar. In addition, it is shown that there is an angular separation of bulk and surface signals due to dispersion.

Before it concludes, Chapter 3 discusses how to use the results of this theoretical analysis to guide microscopic calculations. Also discussions are included on how to experimentally search for resonant RISOM signals. The results of this chapter can be seen as guide to be used by other researchers in doing microscopic calculations and experiments using RISOM to study surfaces.

## 2. INTERFACE-SPECIFIC OPTICAL MIXING EMPLOYING GUIDED WAVE GEOMETRY

### 2.1 Introduction

As discussed in Chapter 1, in optical mixing processes typically the surface signal arises from a region only a few layers of molecules thick. The number of surface-residing molecules that interact with the optical fields (generating mixed fields) is relatively small. This is to be compared with those bulk-residing molecules that ordinarily contribute to the mixed fields. Hence, the bulk contribution to the total signal is usually much greater than the surface contribution. By some contrivance of the geometry of the incident waves, the mixed (i.e., generated) waves are to be created such that the surface-to-bulk signal ratio is enhanced. With this condition, the surface region can be investigated for frequency response, symmetry information, and perhaps time-wise transitory behaviour. This, in turn, aids the researcher in determining the stoichiometry, 'binding' mechanisms, possible migration times, and even (chemical) reaction properties of the surface region. Recall that the third order susceptibility is responsible for four-wave mixing (FWM). Since the third order susceptibility contains more information about the symmetry of the surface-residing species (than the second order susceptibility), FWM is traditionally chosen as the tool to study surface stoichiometry. Laboratory experiments utilizing FWM are not trivial; the probability of generated mixed light is low. Higher order mixing processes are even more illusive and are not commonly employed. When each of the waves in a FWM process are nearly equal in frequency, dispersion plays a minor role. Dispersion-related phase-mismatch can, therefore, be a minimal effect. CARS-like FWM is the process that best satisfies the criterion that all the waves be of nearly the same frequency. Hence, this chapter will focus on designing ISOM experiments utilizing CARS-like FWM.

How can a surface be studied using CARS-like FWM if the surface-to-bulk signal ratio is minuscule? The only way to improve this ratio for any given surface is to inhibit the mixing process in the bulk region and hope that the surface signal will be attenuated to a lesser degree. One way to do this is to take advantage of any dissimilarities in the angular dependence of FWM in the bulk and surface regions. Another approach is to manipulate the phase-mismatch, and hence the generated (output) intensity, such that it is large in the bulk, but small in the surface region. These approaches are investigated in Chapter 3. Each of these approaches rely on differences of the optical mixing properties of the bulk and surface. Yet another approach attempts to engineer a structure in which the mixing process is zero in the bulk due to an interference condition among the applied waves. A non-symmetric planar waveguide can be used to establish just such an interference condition.

This chapter takes the approach of manipulating the applied fields to achieve an interference condition everywhere except the surface region. Thus, fields are only generated in the surface region; this enhances the surface-to-bulk signal ratio. This is done in the confines of an asymmetric planar waveguide. This work refers to ISOM in waveguide geometry as waveguide ISOM (WISOM). The goal of this chapter is to investigate the experimentally relevant parameter space of WISOM. Phase-matching is considered in the waveguide geometry as a way to increase the surface signal. The *ad hoc* searching for phase-matching conditions experimentally is not practical. Hence the goal of exploring the entire parameter space. No definitive body of information or ‘data base’ currently exists on waveguide parameters for WISOM experiments; this work partially remedies that deficiency.

This chapter begins with an overview of CARS-like FWM processes, the details of asymmetric waveguide structures, and the aspects of WISOM experiments. This is done in Section 2.2. The next section, Section 2.3, is devoted to reviewing the work of previous experimental researches in WISOM and gives some justification as to the

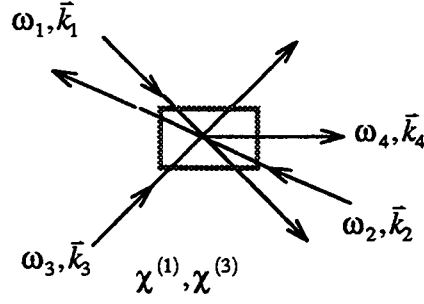


direction taken by this research. Section 2.4 is devoted to the theoretical development of WISOM. In this section, the expressions used to quantify the degree of surface-specificity are derived. These quantifiers include the ‘direct ratio’ and the ‘cross ratio’. The next section of this chapter reports the results of the numerical calculations that were used to evaluate the direct and cross ratios. This section begins with a discussion of what parameters are experimentally relevant and the range of values used in the calculations. It then continues with a detailed description of the numerical analysis of WISOM parameter space. Although all the results cannot be shown in this document, the most relevant data are shown. All the data and the software used generated the data, however, are reproduced on the CD-ROM enclosed at the end of this document. After some initial conclusions are drawn, Section 2.6 continues the discussion of WISOM by considering phase-matching of Gaussian beams in a waveguide. Since phase-matching in a waveguide is advantageous and multiple couplers are required, Section 2.7 considers approaches to fabricating integrated waveguide couplers. Couplers such as achromatic prisms, linear and ring gratings, and photorefractive-based couplers are reviewed.

## 2.2 Overview of CARS-Like FWM and WISOM Experiments

In FWM, four waves with frequencies  $\omega_1$ ,  $\omega_2$ ,  $\omega_3$ , and  $\omega_4$  are involved. Often, the wavevectors are used to describe the process. The wavevectors describing this process are labeled  $\vec{k}_1$ ,  $\vec{k}_2$ ,  $\vec{k}_3$ , and  $\vec{k}_4$ . CARS-like FWM experiments are done in a material having a non-zero third order susceptibility and are typified by  $\omega_4 = \omega_1 - \omega_2 + \omega_3$ , where  $\omega_1 = \omega_3$ ; in terms of the wavevectors,  $\vec{k}_4 = \vec{k}_1 - \vec{k}_2 + \vec{k}_3$ . The geometry of a such an experiment is shown in Figure 2.1. Note that it is convenient to have both waves carried in the same beam by using collinear wavevectors  $\vec{k}_1$  and  $\vec{k}_3$ . This makes the experimental apparatus less complicated.

There is no reason to limit these discussions to a particular FWM process. A general convention can be made to include all processes. The convention adopted is that all



$$\omega_4 = \omega_1 - \omega_2 + \omega_3$$

Figure 2.1: A typical FWM experiment.

annihilated waves have positive wavevectors and frequencies, and all stimulated waves have negative wavevectors and frequencies. Furthermore, an electric field with negative frequency argument is the complex conjugate of the field:  $\vec{E}(-\omega) = \vec{E}^*(\omega)$ . Hence, the observed generated wave's frequency for all FWM processes could be written as  $-\omega_4 = \omega_1 + \omega_2 + \omega_3$ . If the medium has dispersion, the wavevectors of the 'incident' waves are frequency dependent. The induced polarization is written as

$$\vec{P}^{(3)}(\omega_4) = \chi^{(3)}(\omega_4; \omega_1, \omega_2, \omega_3) \vec{E}(\omega_1) \vec{E}(\omega_2) \vec{E}(\omega_3) e^{i[(\vec{k}_1 + \vec{k}_2 + \vec{k}_3) \cdot \vec{r} - \omega_4 t]}. \quad (2.1)$$

The generated wave is characterized by

$$\vec{E}(\vec{r}, t) \sim \vec{E}(\omega_4) e^{i[(\vec{k}_1(\omega_1) + \vec{k}_2(\omega_2) + \vec{k}_3(\omega_3) + \vec{k}_4(\omega_4)) \cdot \vec{r} - \omega_4 t]}. \quad (2.2)$$

Clearly the phase of the source polarization is generally not equal to the phase of the generated wave. That is, the phase-mismatch, measured by the size of

$(\vec{k}_1 + \vec{k}_2 + \vec{k}_3 + \vec{k}_4) \cdot \vec{r}$ , is not generally zero. When the media have dispersion, the phase-mismatch can only be zero if the directions of the wavevectors  $\vec{k}_1$ ,  $\vec{k}_2$ , and  $\vec{k}_3$ , are appropriately adjusted. Each wavevector is described by its spherical coordinates:  $\vec{k}_i = [|\vec{k}_i|, \theta_i, \phi_i]$ , ( $i = 1, 4$ ). As will be discussed, there will only be specific polar angles,  $\theta_i$ , allowed in a waveguide; these correspond to the allowed waveguide modes. Thus, only the azimuthal angles,  $\phi_i$ , allow the experimenter to freely adjust the wavevectors,  $\vec{k}_1$ ,  $\vec{k}_2$ , and  $\vec{k}_3$ , to achieve phase-matching. Due to the physical difficulties, phase-matching in a waveguide has been unattainable by experimentalists. The reader is reminded that when there is no phase-mismatch, the generated signal becomes very large. As the phase-mismatch increases, this signal is attenuated rapidly.

As an example, consider standard CARS-like FWM. Waves of frequency  $\omega_1$  and  $\omega_3$  are annihilated, the wave at  $\omega_2$  is stimulated, and the wave with frequency  $\omega_4$  is generated and subsequently observed. Thus, the sum of all the wavevectors in the process is  $\vec{k}_1 + (-\vec{k}_2) + \vec{k}_3 + (-\vec{k}_4)$  and the frequency of the generated wave is  $\omega_4 = \omega_1 + (-\omega_2) + \omega_3$ . The source polarization would be expressed as

$$\vec{P}^{(3)}(\omega_4) = \chi^{(3)}(\omega_4; \omega_1, \omega_2, \omega_3) \vec{E}(\omega_1) \vec{E}^*(\omega_2) \vec{E}(\omega_3) e^{i[(\vec{k}_1 - \vec{k}_2 + \vec{k}_3) \cdot \vec{r} - \omega_4 t]}. \quad (2.3)$$

The generated wave would, in this case, be characterized by

$$\vec{E}(\vec{r}, t) \sim \vec{E}(\omega_4) e^{i[(\vec{k}_1(\omega_1) - \vec{k}_2(\omega_2) + \vec{k}_3(\omega_3) - \vec{k}_4(\omega_4)) \cdot \vec{r} - \omega_4 t]}. \quad (2.4)$$

A diagram of an asymmetric waveguide typically used for WISOM is pictured in Figure 2.2. The waveguide is constructed of a thin high-index film atop a low-index substrate. Archetypal films have indices between 1.6 and 2.3 and thicknesses,  $h$ , ranging

from 400 to 1000 nm. The substrates are usually 7059 glass or fused silica, which has a lower index of refraction of 1.47. The surface of the film is the transition region between the vacuum and the high-index film. This interstitial region contains unique components that make it different from the film layer. Its thickness is typically on the order of an atomic or molecular diameter, approximately 5 nm or so. Clearly, the surface (or cover) region has different optical properties (both linear and nonlinear) than the film of the substrate. In fact, when molecules are adsorbed to a surface, as is illustrated in Figure 2.3, a whole new set of (nonlinear) optical resonances are added to those of the bulk (i.e., the film and substrate). Often, the resonances of the cover layer of adsorbed species do not significantly overlap the resonances of the waveguiding film. However, the number of species that interact in the nonlinear mixing process for the cover layer is much smaller than those residing in the film. The third order susceptibility of the cover may be larger than the film by a factor of 10 or 10,000; this is because the mixing process accesses a resonance of the cover. Even though the cover has a much larger susceptibility than the film, the film could still contribute more to the total signal due to the much larger number of species that interact with the optical fields. The substrate is usually ignored in this analysis since the evanescent fields do not reach very far into this region. However, it should not be overlooked that the substrate may have its own resonances that the experimenter may have to contend with.

It should be pointed out that the interface between the film and the substrate is likely to have a morphology similar to the film-vacuum interface. A small interstitial region between these two layers may also have unique optical properties. This is something the experimenter wishes to avoid, but it should not be ignored.

Quietly neglected is the fact that the film, since it is to support guided modes, must be an optically transparent dielectric. The restrictions are even more harsh: Not only must the film (i.e., the surface under study) have a high index, it must be a material which can be sputtered, spun-coated, dip-coated, evaporated, or otherwise affixed to the substrate.

The film must be robust enough to survive handling, repeated vacuum cyclings, perhaps thermal cyclings, and the adsorbate molecules. Photoactive and photoreactive materials are usually avoided [28]. These restrictions put severe limits on the types of surfaces that can be studied.

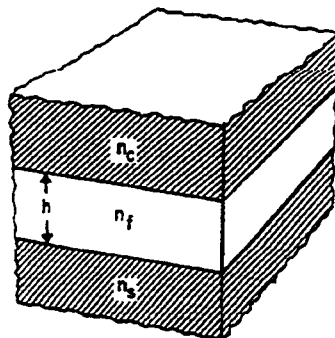


Figure 2.2: Structure of a planar waveguide.

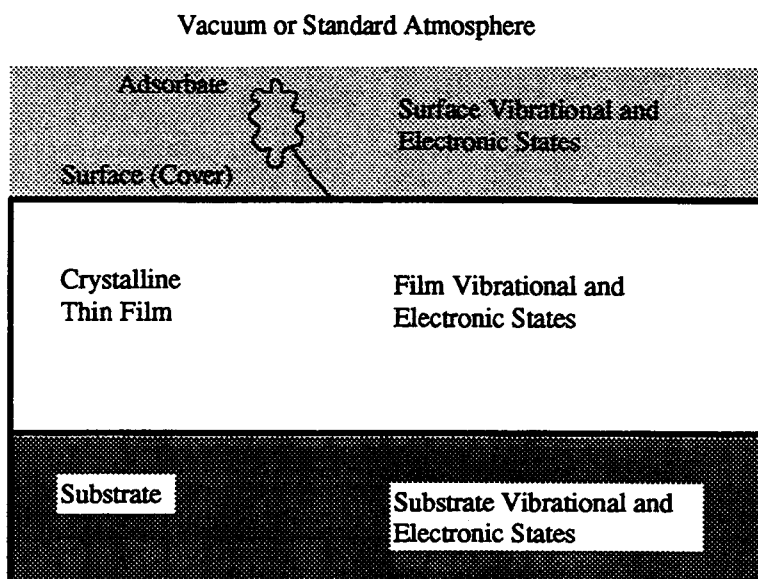


Figure 2.3: A cover layer of adsorbates 'bound' to the crystalline thin film of a planar waveguide.

One impetus for beginning this mapping of waveguide parameters for favourable WISOM conditions was the development of a sputtering technique for laying silicon-oxynitride films,  $\text{SiO}_x\text{N}_y\text{H}_z$  [29]. Using this technique, robust films can be grown on a silicon dioxide substrate. What is unique about this method is that the index of refraction can be selected within the range 1.6 to 2.05 by varying the amount of nitrogen and hydrogen.

A WISOM experiment is illustrated in Figure 2.4. Surface-specificity has been achieved using this planar waveguide geometry [1]. Such a guiding structure consists of the film region of thickness  $h$ , third order nonlinear susceptibility,  $\chi_f$ , and linear index  $n_f$ . The film is grown atop the substrate having nonlinear susceptibility,  $\chi_s$ , and linear index  $n_s$ . Residing above the film is the cover or surface region with nonlinear susceptibility  $\chi_c$  and linear index  $n_c$ . The waves are coupled into the structure using high-index  $\text{TiO}_2$  or  $\text{SrTiO}_3$  prisms. The applied waves are shown entering the left prism in Figure 2.4. The photons with  $k_1$  and  $k_3$  (shown here non-degenerate in wavevector) are annihilated in a CARS-like process;  $k_2$  is used to stimulate the emission of photons at  $\omega_2$ . A wave, with wavevector  $k_4$  is emitted and decoupled from the film by prism on the right. Imagine that the waves shown in Figure 2.4 are propagating in different azimuthal directions. (See, for example, Figure 2.7 or Figure 2.8.) At the location at which the three applied waves cross, the field strengths may be large enough to create the fourth wave. If the susceptibilities (film, cover, and substrate) of this ‘interaction region’ are large, the intensity of fourth wave will be significant. As stated earlier, if the waves are tuned to a resonance of the surface, the ratio of the cover-to-film susceptibilities may be 100. Since the interaction region extends a distance about one vacuum wavelength into the cover region via the evanescent field, all surface species are sampled. Even though the fields may have a large amplitude, the generated fourth wave may not be particularly strong. As a precursor to future discussion, it might be mentioned that in order to achieve a strong signal, the induced nonlinear polarization, (2.1), must be large and the phase-mismatch small. The induced polarization is large if the susceptibility is large and the product of the applied fields (from waves 1

though 3) is sizable. Also, the strength of the mixed wave depends on the size of interaction region; the larger in interaction region, the larger the number of individual species sampled. If the magnitude of the generated (fourth) wave is large, the waves arising from the film part of the interaction region may also be large. Despite the careful minimizing of the fields in the film region, the interaction region in the film (see Figure 2.4) can be significant. The film part of the interaction region (where waves 1, 2, and 3 cross in the film) may be anywhere from  $10^3$  to  $10^5$  times larger the cover portion of the interaction zone. Without arduous attempts to minimize the fields in the film region, the cover-to-film signal ratio, and hence the surface-specificity, can be very poor.

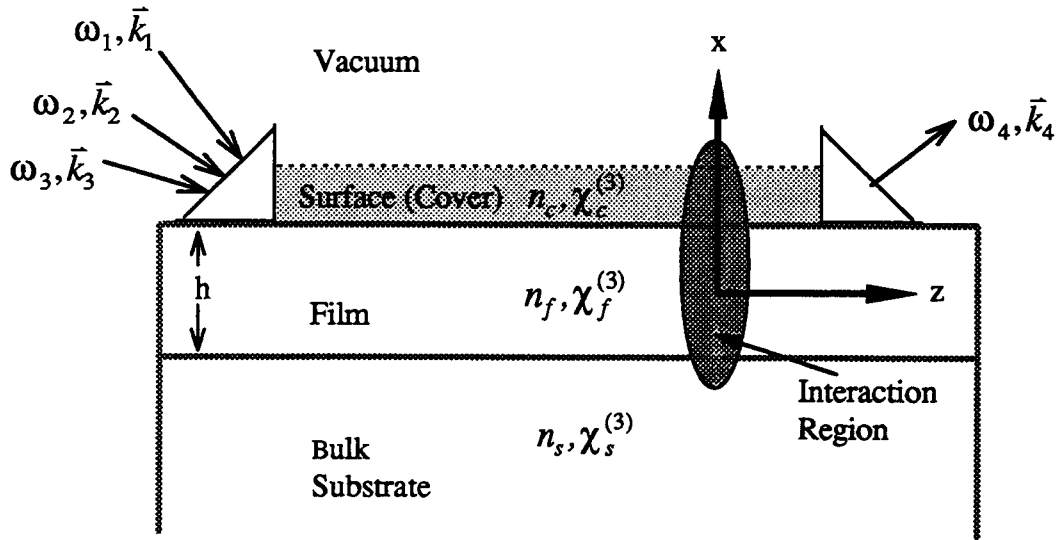


Figure 2.4: Diagram of FWM in a waveguide.

As alluded to, the scheme used to improve the surface-to-film signal ratio is to create a destructive interference condition within the film region. This destructive interference is not fortuitous; it must be carefully engineered. Guided waves have discrete wavevector modes; these modes can be even or odd, not unlike resonance modes on a string. That is, the solution to the electric field distribution for *one* guided wave is either an even or odd

function with respect to  $x$ , the direction parallel to the normal. This is shown in Figure 2.5 where the amplitude of the fields of two even modes (having slightly differing frequencies) and one odd mode are shown in a waveguide structure. The origin of the coordinate system is placed in the center of the film layer as is indicated in Figure 2.4. Note that this plot is oriented differently than Figure 2.4. In Figure 2.4, the  $x$ -direction is vertical, in Figure 2.5 the  $x$ -direction is horizontal, but the  $x$ -direction still represents the direction parallel to the normal of the planar structure. The heavy vertical lines in Figure 2.5 indicate the boundaries of the film region: the right side is the film-cover interface and the left is the film-substrate interface.

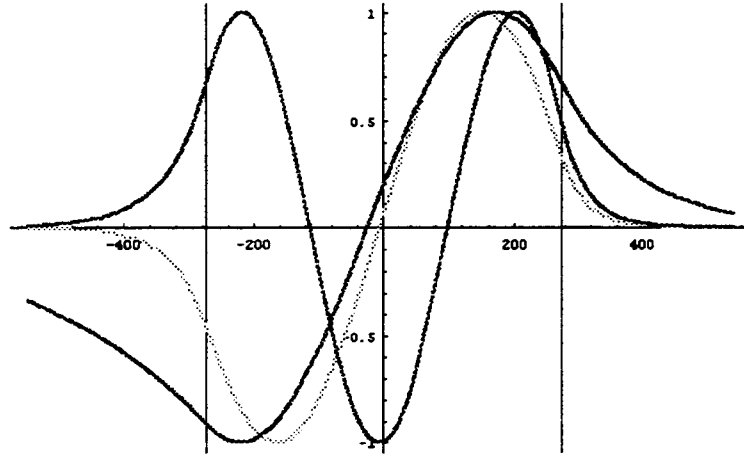


Figure 2.5: The electric field amplitudes of three guided waves versus depth in an asymmetric waveguide.

Recall that the intensity of the fourth wave of the CARS-like process is dependent on the amplitude of the induced nonlinear polarization. The polarization, in turn, is dependent on the product of the fields 1 through 3. Since the induced (fourth) field comes from the *entire* interaction region, the strength of the observed wave is integrated over the interaction region. The integral is a three-dimensional integral over  $x$ -,  $y$ -, and  $z$ -



directions. Clearly, if the product of the three applied fields is antisymmetric with respect to the center of the film, the integral over the film is zero. Hence the mixed wave amplitude from the film is zero. The product of the three field amplitude profiles along the  $x$ -direction shown in Figure 2.5 is shown in Figure 2.6. Here between the two heavy vertical lines, indicating the film region, the integral over this function is (approximately) zero; it is roughly asymmetric about the symmetric limits  $-h/2$  to  $h/2$ . Thus, if two of the waves are coupled to even modes and the third to an odd mode, the product of the fields in the  $x$ -direction is an asymmetric function (about the center of the film layer), the integral over this film vanishes, and the strength of the mixed wave is zero in the film. Note that there are many other combinations of even and odd modes which will (approximately) yield no mixed signal originating from the film. In fact there are 18 such combinations which have this potential (assuming all waves are non-degenerate in wavevector). In the example given above (shown in Figure 2.5), the integral over the cover region is clearly *not* zero. This can be seen by examination of the product of the field amplitudes (Figure 2.6) in the cover region (to the right of the heavy right-hand vertical line); their product is not odd. Thus, there is (approximately) no contribution from the film; the observed mixed wave is generated predominately from the surface region. It might be noted that, from examination of the substrate portion of Figure 2.6, the contribution from the substrate

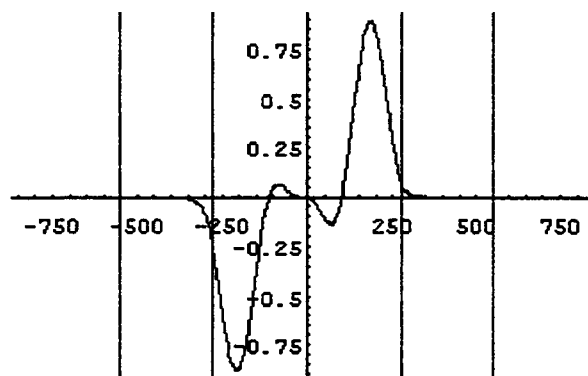


Figure 2.6: The product of electric field amplitudes of three guided waves versus depth in an asymmetric waveguide.

region should not be neglected. Also, it might be pointed out that the integrals in the directions in the plane of the waveguide (y- and z-directions) cannot be zero for a generated wave of finite intensity. Furthermore, the reader is reminded that the phase-mismatch depends of the directions of these applied waves. Hence, the mixed wave may be very weak due to large mismatches even with the favour of all other factors.

The size of the interaction region effects the strength of the mixed wave. The larger the interaction region, the larger the strength. The interaction region depends only on the volume of overlap of the waves. Plane waves are typically used [2] to develop a model for the waveguide, however true laboratory experiments use finite beams. Actual experiments use beams that are Gaussian in intensity profile; these are *focused* into the interaction region. A Gaussian beam experiment is illustrated in Figure 2.7. Here two Gaussian beams (one containing both sources for waves 1 and 3) are focused where they cross. In the region of overlap (i.e., the interaction region) waves of the mixed frequency add together to form a mixed beam. When the beams are not angularly separated by large amounts, the overlap is large as in Figure 2.7. Experimentally, in this collinear beam case, the generated limited by the coherence length of the (pulsed) lasers used. However if the beams are separated by large angles, as in Figure 2.8, the interaction region is smaller. Apparently a larger overlap is desired, or is it? If the case shown in Figure 2.7 causes there to be large phase-mismatches, the generated beam will suffer losses in intensity. One question this research wishes to address is, which case is more favourable? Is a large interaction region with a large mismatch better than the case shown in Figure 2.8, where the process is phase-matched, but the interaction region is small? The reader should be reminded that since only discrete waves are allowed in the waveguide in the vertical direction, the three-dimensional Gaussian nature of a beam coupled to a waveguide is lost. Focused Gaussian profiles remain in the plane of the waveguide; integrations over these dimensions should reflect that. The integration in the waveguide's normal direction is not over a product of three Gaussian profiles. Rather it is the field profile determined by the guided modes in Figure 2.5.

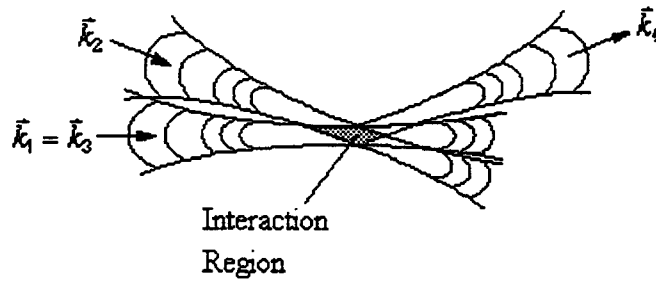


Figure 2.7: Focused nearly collinear Gaussian beams crossing in a nonlinear medium resulting in a large interaction region.

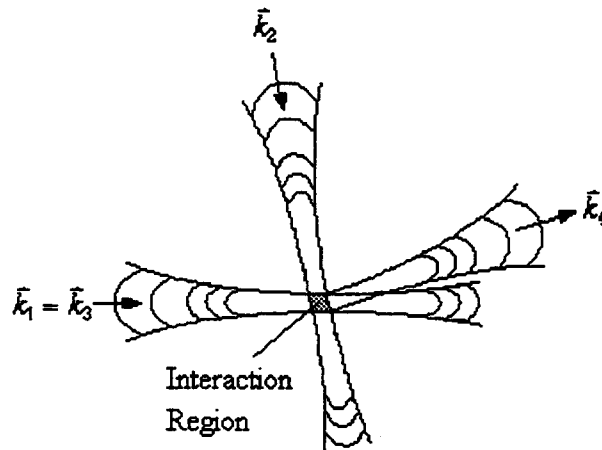


Figure 2.8: Focused Gaussian beams crossing at large angles yielding a small interaction region.

One final subject important to WISOM experiments is the coupling scheme. Commonly, high-index prism couplers are used ([2], [3], and [30]). A  $\text{SrTiO}_3$  or a  $\text{TiO}_2$  prism is pressed against the waveguide such that good optical contact is maintained. Such a scheme is pictured in the left side of Figure 2.9. This shows one prism coupler coupling

three beams (two are degenerate) to an asymmetric waveguide. A second coupler decouples the beam generated in the waveguide. Note that the beams originally coupled into the waveguide are decoupled together with the generated beam. The right side of Figure 2.9 shows two grating couplers used in place of prisms. Grating couplers have been used to couple to planar waveguides in many other applications. Fabrication of linear grating couplers have been reviewed by several researchers [32], [33], [34]. They all are created by standard holographic lithography and etching of the waveguide. This work later discusses fabrication of grating couplers and suggests new grating configurations (other than linear gratings) in an attempt to make phase-matching in a waveguide more attainable. The mode combinations using a single beam containing degenerate frequencies  $\omega_1$  and  $\omega_3$  will be referred to as *three-beam mode combinations*. In the following discussion, the mode combinations are labeled using the convention:  $[v_1, v_2, v_3, v_4]$ , where  $v_1$  is the mode that wave  $k_1$  is coupled to, etc. The three-beam mode combinations are one in which  $v_1 = v_3$ . Mode combinations  $[0,1,0,1]$  and  $[1,1,1,2]$ , for example, are all three-beam mode combinations which may lead to favourable surface-to-bulk signal ratios. This is clear since even functions of the field profile (in the  $x$ -direction) are associated with even modes and vice-versa. Separating the beam carrying  $\omega_1$  and  $\omega_3$  into two beams leads to ‘four-beam mode combinations’. For example, four-beam mode combinations are  $[1,1,0,1]$ ,  $[2,2,0,1]$ , and  $[2,2,1,1]$ .

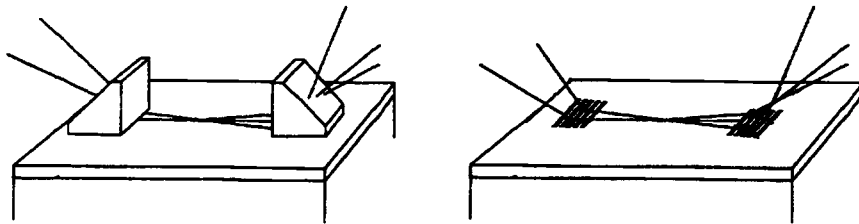


Figure 2.9: Two possible planar waveguide coupling schemes.

### 2.3 Previous Experimental WISOM Work

Four-wave (CARS-like) ISOM performed in dielectric waveguides has been used to provide the surface-specificity sensitive enough to observe fractions of a monolayer of molecules on a surface. CARS was first seen in a waveguide made from a polystyrene film [26]. Surface-specificity was first reported in a  $\text{TiO}_2$  waveguide on a  $\text{Nb}_2\text{O}_5$  substrate [27]. Since then, the WISOM technique has been used to characterize vibrational states of  $\text{O}_2^-$  [1], ethylene [2], and phenol and pyridine [3] on zinc oxide waveguide surfaces. Each of these studies used WSCARS utilizing the interference condition as reviewed in the previous section. This section recalls the important highlights and conclusions of previous experimental WISOM work.

The CARS spectrum of  $\text{O}_2^-$  was investigated on a ZnO surface using a waveguide constructed of a ZnO film sputtered on a silica substrate [1], [31]. The film, with index 2.00, was made by RF sputtering to a thickness of  $600 \pm 30$  nm. Waves were coupled by means of  $\text{SrTiO}_3$  prisms. Two beams, one beam carrying degenerate frequencies  $\omega_1$  and  $\omega_3$ , and one beam carrying  $\omega_2$  were coupled into the waveguide. The beams were focused into the sample via a 300 mm lens. The waveguide was placed in a vacuum chamber and was baked out. As  $\text{O}_2$  was added to the system, frequency scans were taken. With intense beams, a spectrum identified as  $\text{O}_2^-$  was revealed. The surface-to-film signal ratio was reported to be 1:4. Adsorption of ethylene [2] was studied on ZnO surfaces using the same technique and experimental details as the  $\text{O}_2^-$  experiments. A third series of experiments used CARS-WISOM to observe and phenol and pyridine [3] adsorbed on ZnO surfaces. With the mode combination used ([1,1,1,2]), the waveguide was shown to provide surface-specificity over a  $100 \text{ cm}^{-1}$  tuning range. In all three of these experiments, phase-matching was not achievable due to geometrical limitations. Typical angular separation between input beams was 3 to 5 degrees. It was reported that the contribution

in the film was sensitive to the separation angle, indicating that phase-mismatch is important. This report lends justification for the concept that phase-matching in the cover region may enhance the surface-specific signal.

The reader may notice that only a few of the 18 possible guiding modes were used in the experimental work reviewed above. This is partially due to geometrical limitations of the experimental set ups. That is, these experiments all use three-beam mode combinations. Using four-beam mode combinations increases the probability of obtaining surface-specificity. A second realization these experiments provide is that these mode combinations are only surface-specific over a limited frequency range. The original research goal of this work was to experimentally probe the *electronic* states of simple molecules on a surface by means of CARS-like WISOM. The availability of silicon-oxynitride films on fused silica substrates opened exciting possibilities. These waveguides can be fabricated with film indices from 1.6 to 2.1. After an ephemeral theoretical analysis of the WISOM technique for *electronic* states, the experiment appeared to have little promise due to low signal levels. The reported experimental prevailing literature considered only three-beam combinations of waveguide modes to achieve the interference condition necessary for surface-specificity. Furthermore, no researcher has reported phase-matching in the waveguide as a way to increase signal levels. Hence, an extensive investigation was launched to map out all experimentally relevant WISOM parameter space. The motivation being to ascertain if using four-beam waveguide modes and possibly phase-matching could make WISOM a viable technique.

## 2.4 Theoretical Development of WISOM

This theoretical development is aimed at deriving ratios that can be used to quantify the degree of surface-specificity in a WISOM experiment. As it happens, the surface-to-bulk intensity ratio is not the best ratio to investigate due to its divergence when the bulk signal goes to zero. The numerical difficulties in dealing with such divergences can be

avoided by choosing ratios that do not diverge as the parameter space of the waveguiding structure is investigated. In order to map out the parameter space for FWM, the observed field, assumed to have wavevector  $\vec{k}_4$ , must be solved for. The general development done here expresses the field of this observed (plane) wave in terms of the parameters of the incident plane waves, the linear indices, and nonlinear susceptibilities of the cover, film, and substrate. Only transverse electric (TE) modes are considered here.

There is an unfortunate confusion in coordinates used to describe the directions of the wavevectors and fields. For reasons which are apparent in the focused Gaussian beam analysis of WISOM, it is best to describe the plane waves comprising the Fourier decomposition of the beams in a coordinate system with the  $z$ -axis oriented along the direction of the center of the beam. For every beam, there is a separate coordinate system. The waveguide has its own coordinate system which is shown in Figure 2.4. The  $z$ -axis is taken to be along the direction of the first beam, the beam that carries  $\omega_l$ . To best describe the incident and observed plane waves (and beams), the waveguide coordinate system is used. Since the  $z$ -axis is along the propagation direction of the first beam, the usual spherical angles cannot be used to describe the directions of the waves outside the waveguide (before they are coupled or after they are decoupled from the waveguide). Outside the waveguide each wave  $i$  ( $i = 1, \dots, 4$ ) is described by its field  $E_i$  and wavevector  $\vec{k}_i$ . The direction of  $\vec{k}_i$  is defined by  $\vec{k}_i = [|\vec{k}_i|, \vartheta_i, \alpha_i]$ , where the angle  $\vartheta_i$ , is the angle from the  $x$ -axis to the wavevector (the akin to the polar angle) and the angle  $\alpha_i$  from the  $z$ -axis to the wavevector (an *azimuthal* angle). This arrangement is pictured in Figure 2.10.

To quickly refer to the frequencies and angular parameters of the waves, the following convention is adopted: All the frequencies for mode combination  $[v_1, v_2, v_3, v_4]$  are represented by  $[\omega_1, \omega_2, \omega_3, \omega_4]$ , the set of ‘polar’ angles by  $[\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4]$ , and the ‘azimuthal’ direction angles by  $[\alpha_1 \equiv 0, \alpha_2, \alpha_3, \alpha_4]$ . This defines the frequency and directions of the wave *before* being coupled into the waveguide.

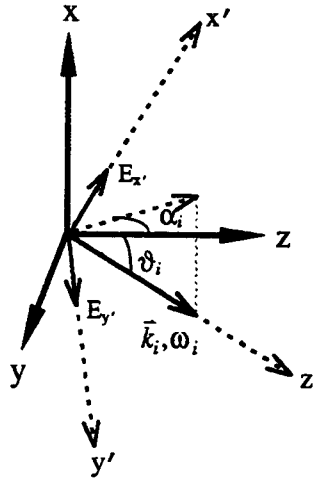


Figure 2.10: Diagram showing the direction angles outside the waveguide structure.

#### 2.4.1 The Wave Equation for Waveguides

To find the contributions to the mixed waves from the cover, film, and substrate, the induced polarizations in each region must be found. This requires that each of the applied fields be known (recall (2.1)). The fields are found by considering Maxwell's equations and applying the boundary conditions at each interface. Usually only plane waves are considered in this solution. The effects of Gaussian beams will be found by making linear superpositions of the plane wave solution; this occurs in Section 2.6. The discussion presented here is detailed in Appendix A. It begins with Maxwell's equations for dielectrics in absence of sources (using Gaussian units) which are:



$$\begin{aligned}\bar{\nabla} \times \bar{E} &= -\frac{1}{c} \partial_t \bar{B}, & \bar{\nabla} \times \bar{H} &= \frac{1}{c} \partial_t \bar{D}, \\ \bar{\nabla} \cdot \bar{D} &= 0, & \bar{\nabla} \cdot \bar{B} &= 0.\end{aligned}$$

In general, all the fields are complex and are functions of space  $(x, y, z)$  and time. The wave equation is derived in the usual way [74]. Making only the assumptions that the index of refraction is time-independent (i.e., ignoring the nonlinear index of refraction, such that  $\partial_t n = 0$ ) and that the waveguide is constructed from dielectric materials, a wave equation for a waveguide can be derived. Since the wave equation for the magnetic field may be easier to solve than is electric field counterpart, both the electric field and magnetic field versions of the wave equation for an inhomogeneous dielectric waveguide, that is, which a continuous variation in  $n$ , are derived in Appendix A. The results are:

$$\begin{aligned}\nabla^2 \bar{E} - \frac{n^2}{c^2} \partial_t^2 \bar{E} \\ = - \frac{2(\bar{E} \cdot \bar{\nabla} n) \bar{\nabla} n + 2(\bar{\nabla} n \cdot \bar{\nabla}) \bar{E} + 2\bar{\nabla} n \times (\bar{\nabla} \times \bar{E}) + n^2 (\bar{E} \cdot \bar{\nabla} n) \bar{\nabla} \left( \frac{1}{n^2} \right)}{n^2}\end{aligned}\tag{2.5}$$

and

$$\nabla^2 \bar{B} - \frac{n^2}{c^2} \partial_t^2 \bar{B} = \frac{2}{n^2} [(\bar{\nabla} n) \times (\bar{\nabla} \times \bar{B})]\tag{2.6}$$

Considering the assumptions made, the wave equations for a waveguide, (2.5) and (2.6), are perfectly general for dielectric waveguides. This includes graded-index and asymmetric waveguides. There are various approaches to solving the above wave equations. Traditionally, solutions to these inhomogeneous differential equations are solved by adding a particular solution to the solution to the homogeneous wave equation. An unorthodox, but effective, approach treats the modes as quantized solutions to the Hamiltonian appropriate to this wave equation and engages perturbation theory. Although the experimental waveguides available [29] for WISOM experiments are all step index waveguides, it is important to review the possibility of inhomogeneous waveguide solutions to the wave equation. The possibility of constructing controlled graded-index waveguides amplifies this import. Furthermore, possible inhomogeneities in the ( $x$ - $z$ ) plane of the waveguide caused by surface roughness or uneven thicknesses may impose the need for considering such inhomogeneities. Finally, much of the work in inhomogeneous waveguides is applicable to inhomogeneous fields. In particular, inhomogeneous fields caused by coupling Gaussian beams to a homogeneous waveguide. Thus, a brief summary of these solutions is presented in the next section.

#### 2.4.2 Review of Solutions to the Wave Equation for Inhomogeneous Waveguides

Traditionally, descriptions of beam propagation in inhomogeneous multimode graded-index waveguides are constructed by employing the scalar Helmholtz wave equation for a monochromatic component of the beam field. (See, for example, [35, 38, 39, 40].) The Helmholtz equation for the general case cannot be solved exactly, and different methods of approximation must be employed for its solution. It might be noted, however that for homogeneous waveguides with focused Gaussian input fields, the solutions can be found by numerical methods. This is the approach the research presented here eventually takes. The multimode graded-index waveguides can usually be characterized by two small parameters:  $a/L \ll 1$  and  $1/(k a) \ll 1$ , where  $L$  represents the longitudinal scale of any

inhomogeneity,  $k$  is the magnitude of the wavevector, and where  $a$  is the transverse size of the waveguide. Notice that the parameter  $a$  represents the effective thickness of the waveguide including the thickness of the film and the penetration lengths of the evanescent fields into the cover and substrate regions. These two small parameters ( $a$  and  $L$ ) make it possible to employ different methods of approximation, which combine peculiarities of both the method of small perturbations and the method of short wavelength asymptotes known from diffraction theory (see, for example, [41, 42]). Theoretical investigations of multimode graded-index waveguides based on geometrical optics use approximate methods of solving the wave equation (see, for example, [43] and references cited there). In this method, the beam is represented as superposition of homogeneous plane waves. The phase of each wave in the superposition is characterized by an optical path length evaluated from the Eikonal equation. The amplitude of each wave in the superposition may be calculated from a system of transport equations. A major advantage of the method of geometrical optics is the simple picture it provides for the characterization of beam propagation. This theoretical method offers an accuracy which is quite sufficient for most practical applications and is adopted by this researcher. The main disadvantage of this approach is the difficulty met in taking into account typical wave effects, such as the behaviour of wave fields in the regions where beams come to a focus and also in waveguides with gain or losses. Numerical methods can be employed to deal with these difficulties. Different modifications of this method have been developed to overcome these problems. Maslov's canonical operator method [44, 45, 46], which is a generalization of the WKB method of quantum theory, is a common modification of the geometrical optics method. The methods of evanescent waves [47] and of complex rays [48, 49] were developed for the description of the fields in regions of caustics, shadows, or in absorbing media. Other methods of constructing solutions to the wave equation make use of localized Gaussian beams [50, 51]. Here local inhomogeneities in the index or the field (as in the case of a focused beam) is represented by an integral over flat-field (i.e., unfocused) Gaussian beams. Another powerful approach developed for characterizing mainly longitudinally inhomogeneous waveguides is the method of cross sections [52, 53,

36]. It employs an expansion of the field in the inhomogeneous waveguide in terms of modes of a (longitudinally homogenous) uniform waveguide whose refractive index profile coincides with an index profile of the inhomogeneous waveguide at the given cross section. This method is developed by Borovikov [54, 55]. It has also been proposed by Popov [56] to expand the field in an inhomogeneous waveguide over some other model functions representing generalizations of the modes of a uniform waveguide.

Unfortunately, employing numerical methods like the finite difference or Fourier transform beam propagation method [57] for solving Helmholtz's equation is very time consuming and does not permit analyzing wave propagation in detail. The latter goal can be achieved solely with the help of exact or approximate analytical techniques. Among the diverse methods for obtaining approximate solutions of the Helmholtz wave equation deserving particular attention is one that relies on obtaining solutions to a parabolic (paraxial) differential equation; a proposal going back to Leontovich and Fock [58]. Many researchers have contributed to the development of this method: [37, 59, 60]. The parabolic equation method is applied in order to treat obstinate problems occurring in the theories of diffraction of the propagation of laser beams in waveguides and in laser components [35, 39, 40, 41, 62, 63]. Different approximate methods based on the perturbation theory for a field component (or its logarithm) have been developed in the framework of the paraxial approximation using the parabolic equation [62]. The most important aspect of the parabolic wave equation, however, is that it has a similar form to the Schrodinger equation [35]. Because of this formal resemblance, the well established quantum-theoretical methods may be used in order to obtain analytical solutions. This is excellently reviewed by Krivoslykov [62]. Krivoslykov devoted his book to the theoretical investigation of transverse and longitudinally inhomogeneous multimode graded-index waveguides. The guiding properties of a graded-index waveguide are brought about by the transverse variation of its refractive index. In the analysis, the assumption was made that all waveguide inhomogeneities have characteristic lengths which are small compared with the scale of wavelength. Krivoslykov discusses the propagation of coherent or partially coherent beams through longitudinally

inhomogeneous active or passive multimode graded-index waveguides. This is done utilizing the formalism of quantum theory. Particularly informative examples include the detailed behaviour of modes in waveguides and the control of beam properties in different waveguide structures. Analytical solutions to the parabolic wave equation are derived and the effects of the inhomogeneities (in refractive index) are treated in standard perturbation theory. Another important feature of the quantum-theoretical approach is the employment of the coherent state representation and its generalizations for the description of beam propagation. In particular, standard coherent states correspond to Gaussian beams with plane wavefronts. The disadvantage to this method is that it is difficult to interpret the results in a simple physical manner.

#### 2.4.3 Plane Wave Solutions to the Step Index Waveguide

In this work the approach is taken that the Helmholtz equation can be divided into three equations, one for each region of the asymmetric step index waveguide. Plane wave solutions are sought for each region and matched at the boundaries between regions. Focused Gaussian beam solutions are considered utilizing a linear superposition of plane wave solutions.

Given the planar geometry, the polarization basis consists of transverse electric (TE), that is the electric field in the plane of the waveguide, and transverse magnetic (TM), the magnetic field is in the plane of the waveguide (along the y-axis).

The experimenter measures the mode angles,  $\vartheta$ , and crossing angles,  $\alpha$ . Define a primed coordinate system that is rotated about the  $y'$ -axis an angle  $-\vartheta$  and then about the  $x'$ -axis by an angle  $-\alpha$  from the unprimed system in

Figure 2.10. The electric field in the waveguide coordinate (the unprimed) system

$$\vec{E} = E_x \hat{x} + E_y \hat{y} + E_z \hat{z}$$

can be represented in the plane wave (the primed) coordinate system by use of transformation angles such that

$$\begin{aligned}\vec{E} = & E_{x'} \cos\vartheta \hat{x} + E_{x'} \sin\vartheta \sin\alpha \hat{y} + E_{x'} \sin\vartheta \cos\alpha \hat{z} + \\ & E_{y'} \cos\alpha \hat{y} - E_{y'} \sin\alpha \hat{z}\end{aligned}$$

In terms of the TE polarization in the unprimed frame, the fields are described by

$$\vec{E}_{TE} = E_{y'} \cos\alpha \hat{y} - E_{y'} \sin\alpha \hat{z} \quad (2.7)$$

and in terms of the TM polarization by

$$\vec{E}_{TM} = E_{x'} \cos\vartheta \hat{x} + E_{x'} \sin\vartheta \sin\alpha \hat{y} + E_{x'} \sin\vartheta \cos\alpha \hat{z}. \quad (2.8)$$

With expressions (2.7) and (2.8) and the solutions to the wave equations, any combination of waves (and Gaussian beams) can be described. Referring to Figure 2.10, the wavevector in the unprimed frame is

$$\vec{k} = -k_{z'} \sin\vartheta \hat{x} + k_{z'} \cos\vartheta \sin\alpha \hat{y} + k_{z'} \cos\vartheta \cos\alpha \hat{z}, \quad (2.9)$$

with

$$|\vec{k}| = k = nk_o = n \frac{\omega}{c} = k_{z'}.$$

Here  $k_o$  is related to the vacuum wavelength,  $\lambda_o$ , by  $k_o = 2\pi/\lambda_o$ . This analysis allows for a relation between the fields strength of the waves (or beams) outside and inside the waveguide to be derived.

An example of an asymmetric step index waveguide is shown in Figure 2.4. There, the index of refraction changes along the perpendicular, or  $x$ -direction, in a discontinuous way. Recall that the coordinate system of the waveguide places the origin in the center of the waveguide. If the film thickness is  $h$ , then the cover begins at  $h/2$  and the substrate-film interface is found at  $-h/2$ . Thus, the index dependence on  $x$  could be written as:

$$n = n(x) = \begin{cases} n_c & \text{for } x > \frac{h}{2} \\ n_f & \text{for } |x| < \frac{h}{2} \\ n_s & \text{for } x < -\frac{h}{2}, \end{cases} \quad (2.10)$$

where  $n_c$  is the index of the cover,  $n_f$  is the of the film, and  $n_s$  is the index of the substrate.

Similarly, the nonlinear susceptibility can be expressed in a step-wise fashion as

$$\chi^{(3)}(x) = \begin{cases} 0 & \text{for } x > \frac{h}{2} + \text{one monolayer} \\ \chi_c & \text{for } \frac{h}{2} < x \leq \frac{h}{2} + \text{one monolayer} \\ \chi_f & \text{for } |x| < \frac{h}{2} \\ \chi_s & \text{for } x < -\frac{h}{2}. \end{cases} \quad (2.11)$$

Solutions to the Helmholtz equation for the three regions are then found. (The solution for the vacuum region is well known.) Boundary conditions are applied to match the solutions in the usual way. Solutions are found by assuming the initial condition that infinite plane waves with a given polarization are impinging at the film-cover boundary of the waveguide. In this situation the electric field at the cover-film boundary has the form

$$\vec{E}(x, y, z, t) = E_y \hat{y} e^{i(\vec{k} \cdot \vec{r} - \omega t)}, \quad (2.12)$$

where

$$\vec{k} \cdot \vec{z} = k_0 n_f z \sin \theta. \quad (2.13)$$

In a ray picture, the angle  $\theta$  is the angle the rays (in the film layer) reflect off the surface-film boundary. The experimenter measures an incident angle  $\vartheta$  outside the waveguide as discussed in (2.9). For TE modes, the electric field amplitudes in three regions of the waveguide (including their phases) are found to be

$$E_y = \begin{cases} \left[ \frac{(n_f^2 - N_v^2)}{(n_f^2 - n_c^2)} \right]^{\frac{1}{2}} e^{-K_c(x-h/2)} e^{i\beta_v z} e^{-i\omega t} & \text{for } x > h/2 \\ & \text{(i.e., in the monolayer)} \\ \cos(K_f x - \phi_v) e^{-i\beta_v z} e^{-i\omega t} & \text{for } |x| < h/2 \\ \left[ \frac{(n_f^2 - N_v^2)}{(n_f^2 - n_s^2)} \right]^{\frac{1}{2}} e^{+K_s(h/2+x)} e^{i\beta_v z} e^{-i\omega t} & \text{for } x < -h/2, \end{cases} \quad (2.14)$$

with the following definitions:  $\beta_v$  is an effective wavevector for the plane wave traveling in the film. It is defined as

$$\beta_v \equiv k n_f \sin \theta_v. \quad (2.15)$$

This effective wavevector is mode dependent as is indicated by (2.18). The mode dependent parameters  $N_v$ ,  $\phi_v$ ,  $K_c$ ,  $K_f$ ,  $K_s$  are defined as



$$N_v \equiv \beta_v / k, \quad (2.16)$$

$$\phi_v \equiv K_f \frac{h}{2} - \phi_c = -K_f \frac{h}{2} + \phi_s, \quad (2.17)$$

$$v\pi = K_f h - \phi_c - \phi_s, \quad (2.18)$$

$$K_c^2 \equiv k^2 n_c^2 - \beta_v^2 \text{ for the cover,} \quad (2.19)$$

$$K_f^2 \equiv -k^2 n_f^2 + \beta_v^2 \text{ for the film, and} \quad (2.20)$$

$$K_s^2 \equiv k^2 n_s^2 - \beta_v^2 \text{ for the substrate.} \quad (2.21)$$

The additional mode dependent parameters which were used in defining the above parameters are

$$\phi_s \equiv \tan^{-1} \left( \frac{K_s}{K_f} \right) \quad (2.22)$$

and

$$\phi_c \equiv \tan^{-1} \left( \frac{K_c}{K_f} \right). \quad (2.23)$$

To find the allowed modes, that is the angles  $\theta_v$  (or  $\vartheta_v$  as measured outside the waveguide), a transcendental equation must be solved:

$$\tan \left( \left[ K_f \right]_v h - v\pi \right) = \left[ K_f \frac{K_c + K_s}{K_f^2 - K_c K_s} \right]_v. \quad (2.24)$$

These final expressions (2.14) through (2.24) describe the field amplitude and phase of a guided wave of frequency  $\omega$  in the three regions of the asymmetric waveguide. A diagram of the field strength versus the transverse direction (i.e., the  $x$ -direction) is shown Figure 2.5 for three guided modes. In Figure 2.5, two of the waves are in the same (odd) mode, but have differing frequencies, the other wave is in an even mode ( $v = 2$ ) and is approximately symmetric about the origin (the center of the waveguide). Looking at (2.14), in the film there is a traveling wave with an effective wavevector that depends on the discrete mode the wave is guided in. In the cover and substrate regions, there is an exponentially decaying evanescent field. The depth to which these fields penetrate the region also is mode dependent. Expression (2.14) describes a guided wave of a single frequency. To accommodate four-wave mixing experiments in a waveguide, four guided waves are required. The next section discusses the four wave mixing process in a waveguide, determines the generated intensity, and derives the ratios used to quantify the surface (cover) specificity.

#### 2.4.4 The Intensity, Phase-Mismatch, and Quantitative Analysis Tools for WISOM

Four wave mixing in asymmetric waveguides can be surface-specific if the induced polarization due to the incident waves is (nearly) zero in the film and non-zero in the cover region. In this treatment, the contribution to the field from the cover and film (including substrate) regions are found separately. To quantify the surface-specificity, two ratios are introduced which are sensitive to the cover-to-film signal ratio. Since the cover-to-film field amplitude ratio diverges as the film contribution goes to zero, it is not convenient as a quantifier of surface-specificity. Moreover, this ratio does not show effects of differences in phase of the cover and film contributions. A more useful set of interpretive tools are needed, hence two peculiar ratios are introduced. One is sensitive to the amplitude of the cover contribution to the overall signal, while the other is sensitive to both the amplitude and phase of the cover contribution in relation to the entire

contribution. Notice that the entire contribution includes the contributions from the cover, film, and substrate regions. It is also important to know the phase-mismatch between the induced polarization and the generated (mixed) wave. Only infinite plane waves are considered in this initial treatment. In determining the intensity of the generated signal, it is realized that there are four separate waves in the FWM process. After coupling into the waveguide from the vacuum these incident waves (or before decoupling these exiting waves from the waveguide) have wavevectors

$$k_1 \equiv n(\omega_1) \frac{\omega_1}{c}, \quad k_2 \equiv n(\omega_2) \frac{\omega_2}{c}, \quad k_3 \equiv n(\omega_3) \frac{\omega_3}{c}, \quad \text{and} \quad k_4 \equiv n(\omega_4) \frac{\omega_4}{c}. \quad (2.25)$$

Recall that stimulated or observed waves have negative frequency by the convention used here. In a CARS-like process, the first three of these waves mix and produce the fourth. Inside the waveguide, they mix over a region called the *interaction region*. The phase-mismatch parameter between the polarization and the generated fourth wave is, in general, given by

$$\Phi(x, y, z) \equiv \Delta \vec{k} \cdot \vec{r} \quad (2.26)$$

with

$$\Delta \vec{k} \equiv (\vec{k}_1 + \vec{k}_2 + \vec{k}_3) + \vec{k}_4, \quad (2.27)$$

where  $\vec{k}_4$  is the wavevector of the observed wave. The convention used is that positive wavevectors (and frequencies) indicate annihilation, negative wavevectors (and frequencies) indicate creation of photons in the OM process. It is understood that since  $\vec{k}_4$  is observed, this wavevector is negative as is  $\vec{k}_2$  for a CARS-like process. The phase-mismatch parameter for the possible guided mode combination  $\mathbf{v} = [v_1, v_2, v_3, v_4]$  in the

asymmetric waveguide is

$$\Phi = \Delta \vec{k}_v \cdot \hat{z} \frac{\ell}{2}, \quad (2.28)$$

where

$$\begin{aligned} \Delta \vec{k}_v \cdot \hat{z} = & k_1 n(\vec{r}) \sin \theta_{v_1} \cos \alpha_{k_1} + \\ & k_2 n(\vec{r}) \sin \theta_{v_2} \cos \alpha_{k_2} + \\ & k_3 n(\vec{r}) \sin \theta_{v_3} \cos \alpha_{k_3} + \\ & k_4 n(\vec{r}) \sin \theta_{v_4} \cos \alpha_{k_4} \end{aligned} \quad (2.29)$$

and  $\ell$  is the interaction length. The interaction length represents the characteristic length over which the three incident waves overlap. Of course, for plane waves  $\ell$  is infinite, but plane waves are not used in practice. Thus,  $\ell$  is approximated by some experimentally reasonable value. Recall the angles  $\alpha$  are the angles in the plane of the waveguide that each wave propagates at. The famous result [65] for the intensity of the fourth wave generated in a FWM process appears as

$$I_{\omega_4} = \left| \int \int \int_{\substack{\text{all space} \\ \text{all time}}} \vec{E}_4 \cdot \chi^{(3)} : \vec{E}_1 \vec{E}_2 \vec{E}_3 \frac{\sin \Phi}{\Phi} dx dy dz dt \right|^2. \quad (2.30)$$

Since pulses are used to obtain large fields, the time integration is done over one pulse. However, in practice the time integration could be replaced by a time average over one period. Simplification of (2.30) can be made by replacing the electric fields with the form

$$E_{y_i}(x) = f_i(x) E_{y_i}$$

which are the solutions to the Helmholtz equation for a step index waveguide given in (2.14). The subscripts  $i$  represent the wave index ( $i = 1, \dots, 4$ ). The intensity of the mixed wave under this substitution becomes

$$I(\omega_4) = \left( \frac{2\pi}{k_4} \right)^2 \left( \frac{\omega_4}{c} \right)^4 I_{\omega_1} I_{\omega_2} I_{\omega_3} \times \left| \int_{-\infty}^{\infty} \hat{\mathbf{e}}_4 \cdot \chi^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \right|^2, \quad (2.31)$$

where  $I_{\omega_1}$ ,  $I_{\omega_2}$ , and  $I_{\omega_3}$  are the maximum amplitudes of the incident waves at  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ , respectively. The interaction length  $\ell$  must be experimentally estimated. Alternatively, the interaction length can be replaced by the maximum interaction length determined by the coherence length  $\ell_c$ . This is the length for which  $\Phi = \pi$  (see (2.28)). Since the waves are considered infinite plane waves, the integral represents an intensity per unit area. Hence, the integration is performed only over the dimension transverse to the waveguide. The integral can be separated into two parts: the integral over the cover region and the integral over the film and substrate regions. It is typical in an experiment to adjust the frequencies of the input waves to incur a resonance in the cover medium and not the film. Hence, the nonlinear susceptibility for the cover is commonly 10 to 10,000 times larger than the film nonlinear susceptibility. This is reflected in the assumed form for the susceptibility:

$$\chi^{(3)}(x) = \begin{cases} 100 & h/2 < x < h/2 + \text{monolayer} \\ 1 & \text{elsewhere.} \end{cases} \quad (2.32)$$

Thus, (2.31) can be rewritten as

$$\begin{aligned}
I(\omega_4) = & \left( \frac{2\pi}{k_4} \right)^2 \left( \frac{\omega_4}{c} \right)^4 I_{\omega_1} I_{\omega_2} I_{\omega_3} \\
& \times \left| \int_{-\infty}^{-h/2} \hat{\mathbf{e}}_4 \cdot \chi_s^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \right. \\
& + \int_{-h/2}^{+h/2} \hat{\mathbf{e}}_4 \cdot \chi_f^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \\
& \left. + \int_{+h/2}^{+\infty} \hat{\mathbf{e}}_4 \cdot \chi_c^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \right|^2.
\end{aligned}
\tag{2.33}$$

The first term inside the absolute square in (2.33) represents the contribution to the total intensity from the evanescent fields that penetrate and mix in the substrate region. The second term describes the contribution from the film layer. The last term represents the contribution to the intensity generated by the fields mixing in the cover region. Clearly, because the integrals are evaluated and then the absolute square is taken, the contributions from each region of the waveguide are inseparable. Simply taking the ratio of the cover-to-film field amplitudes is not a good quantifier of surface-specificity compared to what is experimentally measured. The interpretative tools that best describe surface-specificity are the direct ratio (DR) and the cross ratio (CR). The DR is an indicator of the surface-specificity; it is closely related to the surface-to-bulk intensity ratio. However, the CR is a better indicator of the signal-to-noise ratio. The DR is the ratio of the cover contribution to the intensity versus the total intensity less the cross terms. The CR is the ratio of the cross terms of the absolute square versus the total intensity less the cross terms. To better understand these definitions, consider that the first two terms in (2.33) are the

‘background’ portion,  $E_B$ , of the generated (fourth) field and the last term is the cover-specific anecdotal field,  $E_A$ :

$$E_A \equiv \int_{+h/2}^{+\infty} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_c^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \quad (2.34)$$

and

$$\begin{aligned} E_B \equiv & \int_{-\infty}^{-h/2} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_s^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \\ & + \int_{-h/2}^{+h/2} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_f^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx . \end{aligned} \quad (2.35)$$

Using these definitions, the total intensity is proportional to

$$I(\omega_4) \sim |E_A + E_B|^2 = |E_A|^2 + |E_B|^2 + E_A^* E_B + E_A E_B^*. \quad (2.36)$$

The DR can be easily defined as

$$DR \equiv \frac{|E_A|^2}{|E_A|^2 + |E_B|^2}. \quad (2.37)$$

Finally, the CR can be seen to be

$$CR \equiv \frac{E_A^* E_B + E_A E_B^*}{|E_B|^2}. \quad (2.38)$$

The DR is a better interpretive tool than the cover-to-film field amplitude ratio since the DR never diverges. Also, notice that the denominator of the DR does not include the cross terms that the total intensity does (compare this denominator with (2.36)). If the fields from the surface and film are anti-parallel and of the same magnitude, the total intensity is zero. Again, the DR does not diverge even under these conditions. Hence, the direct ratio is an appropriate tool to use when the phases of the surface and film fields are approximately equal. The DR was chosen so that it is bounded between zero and one. Clearly, a value of one indicates excellent surface specificity. The DR is, however, not sensitive to the phases of the fields. Thus, the cross ratio is used when the phases of the fields are important. The CR examines only the cross terms in the total intensity. For fields that are purely real, the CR measures the degree to which the cover-generated field adds or subtracts from the field generated from the bulk. The CR is not bounded as the DR is. A large positive value indicates parallel fields where the magnitude of the surface is much greater than the bulk. Conditions yielding a large negative CR indicate the surface field is of opposite sign and is much larger than the bulk-generated field; thus the total intensity is less than the bulk signal solely. For effective WISOM it is required that the surface resonance signal overpower the bulk resonance signal. To obtain favourable DRs and CRs, the indices of refraction of film and substrate, thickness of film, and mode combinations must be adjusted to yield the interference condition in the film. Furthermore, the field frequencies must be tuned, especially over those frequencies that yield cover resonances.

## **2.5 Introduction to the Numerical Analysis of WISOM**

An extensive computational investigation was launched to map out all experimentally relevant parameter space in order to locate situations that yield favourable WISOM conditions. This numerical analysis is, in part, to be used as a data base for other researches. No such data base has been made available in the literature. The analysis



encompasses a large range of experimentally important waveguide parameters. As explained in the introduction to this chapter, only three-beam mode combinations have been investigated by previous researchers. Recall that these combinations are those in which the first and third waves are degenerate, that is, when  $\bar{k}_1 = \bar{k}_3$  and the waves are collinear in the same guided mode. One goal of the analysis is to investigate the four-beam mode combinations where the degeneracy in the first and third waves is broken. It was anticipated that the use of four-beam mode combinations would allow for phase-matching inside the waveguide; a circumstance never reported in the literature. With the potential improvements in surface-specificity gained by these situations, WISOM might be a more attractive technique.

The analysis takes the approach of first picking the index of the film and its thickness. The range of waveguide parameters are chosen within the confines of what are realizable in actual waveguides. The substrate index is fixed; it is considered to be that of fused silica. Likewise, the cover index remains fixed throughout the analysis. The nonlinear susceptibilities are also considered fixed, however the susceptibilities of the cover is considered much larger than the film or substrate susceptibilities. With assumed waveguide index profiles, dimensions, and the FWM frequencies; numerical solutions are found for the generated electric field via (2.14) through (2.24). The numerical solutions are found by a Mathematica (V. 2.2) notebook as described in Appendix A. The azimuthal angles for phase-matching are found when the condition is allowed. Finally, from these fields, the direct ratio and cross ratio are tabulated. For each set of physical waveguide parameters, the frequency and mode combinations are varied. The result is a series of shaded plots indicating waveguide thickness on one axis, index of the films on another, and the shading representing the values of the DR or CR. These plots were generated by Spyglass Transform (V. 3.0) from ASCII text files supplied by the Mathematica notebook *cars1.ma* under the *wscars* directory. This raw data generated directly from this Mathematica notebook can be accessed in the form of text files *rawm0a.xls* and *rawm0b.xls*, enclosed in the CD-ROM in the attachments section of this

document. These are the unmodified files in comma delimited ASCII format. These files were modified by removing all spaces and saved as *carsm0a.xl* and *carsm0b.xl* under that same format. Using Excel (V3.0) macro *cars0m1.xlm*, these files are converted to Excel files, separated by mode combination and files suitable for import into Spyglass Transform are created. Detailed documentation is enclosed in this Excel macro file. The file naming conventions for the mode combination data are given in the notebook *cars1.ma* where they are generated. Furthermore, the raw data files are concatenated by three Mathematica notebooks *carsanal.ma*, *carsanal2.ma*, and *carsanal3.ma*. This software sorts the results in various ways for easy access. An entire summary of the results of the calculations can be found in the Microsoft Excel files *cars0a.xls*, *cars0b.xls*, and *cars0c.xls*. Text versions (delimited by tabs) are also available on the enclosed CD-ROM. Two more files contain abbreviated results sorted as to emphasize the effects of mode combination (file *carsmd.xls*) and the phase-matching information (file *carspb.xls*). Spyglass Transform accepts text files (created by macro *cars0m1.xlm*) which are sorted into directories by mode combination under the *wscars\cars0m\_fi* directory and, via the Spyglass macros *create\_DR\_animations* and *create\_CR\_animations*, are converted into false colour (or gray scale) plots. All Spyglass macros have been exported to ASCII files and are available under directory *wscars\spyglas\_* on the CD-ROM. For each mode combination, a series of such plots were generated as a function of frequency of the applied waves. The exact details of the ranges of the parameters used, the key to interpreting the shaded plots, and a review of useful patterns in the results are discussed in the next sections.

### 2.5.1 Common CARS-Like WISOM Parameters

Planar waveguides used for WISOM work are limited by the availability of appropriate media. As far as the numerical analysis is concerned, the limitations placed on waveguides are those placed on the index of the film layer such that they can be physically constructed. This restricts the range of indices of the film from 1.60 to 2.10. In the numerical

calculations, steps in the film index of 0.05 were used. Thus, the film index,  $n_f$ , starts at 1.60 and is increased in steps of 0.05. The index of the film, in turn, places limits on the thickness, since there is a minimum thickness for a waveguide to support a single mode. Modal separation can be achieved by using waveguides with very different mode coupling angles. As the number of supported modes for a given waveguide increases, the mode coupling angle separation (between successive mode coupling angles) decreases. Since modal separation is useful for the experimenter, the number of supported modes is kept low. This enhances modal separation. No more than six or seven supported modes per waveguide are desirable. All variables told, the thickness range of the film that is experimentally useful is 600 to 950 nm. Hence, numerical calculations begin with film thicknesses of 600 nm and increase in steps of 50 nm. The cover index and (the fused silica) substrate indices were taken as the vacuum and 1.47, respectively. The cover index, when studying a sub-monolayer on the film surface is expected to be nearly that of the vacuum. The frequencies for the applied waves, those described by  $k_1$ ,  $k_2$ , and  $k_3$ , are chosen with values appropriate for CARS (vibrational resonances), however the software developed is flexible enough to deal with electronic resonances. The frequencies are selected with  $\lambda_1 = \lambda_3 = 560$  nm. The second wave, described by  $k_2$ , is picked by means of the difference in wavenumbers between the second and first waves. The energy gap between intermediate states is operative variable. Hence, the difference frequency,  $\Delta$ , is defined by  $\Delta \equiv \omega_1 - \omega_2$ . The equivalent range of the difference frequency (also called the tuning parameter) is from 1000 to 3000 wavenumbers (roughly 580 to 670 nm) in steps of  $200 \text{ cm}^{-1}$ . All possible mode combinations of these four waves that might lead to an interface-specific signal were considered. The mode convention is  $[v_1, v_2, v_3, v_4]$  for guided waves  $[k_1, k_2, k_3, k_4]$ . The set of ‘polar’ angles determined by the mode combination are referred to by  $[\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4]$ ; they are numerically calculated via (2.24). The goal of these calculations is to identify experimentally compatible parameters for CARS-like four-wave mixing inside a waveguide such that surface sensitivity is maximized. In addition, the values of the ‘azimuthal’ direction angles for each mode combination,  $[\alpha_1 \equiv 0, \alpha_2, \alpha_3, \alpha_4]$ , which give phase-matching are found (when phase-

matching is possible). These angles are found by minimizing the phase-mismatch parameter,  $\Phi$ , using expressions (2.28) and (2.29). The experiment is (mechanically) simplified if all angles  $\alpha$  are between -10 and 10 degrees or -170 and 170 degrees. Thus, situations where the frequencies and mode combinations yield phase-matching for small azimuthal angles are particularly sought. To compare with the phase-matched case, the 'large-overlap experiment' is analyzed (for three-beam mode combinations only). The large-overlap experiment is done by using degenerate wavevectors  $k_1$  and  $k_3$  and making the angular separation between  $k_2$  and  $k_1$  three degrees or less. Although the large-overlap experiment has a large phase-mismatch, the interaction region is large. One aim of these calculations was to determine if the phase-matched situations with their small overlap or the large-overlap experiment yield better cover-to-film signal ratio.

The values of  $\chi^{(3)}$  for the various regions of substrate, film, and vacuum were taken as unity, except over the cover region 0.5 nm above the film. This simulates the effect of a one monolayer thick absorbate on the surface of the film that is in resonance with the mixing process. Here  $\chi^{(3)} = 100$  for the monolayer. Hence, in (2.11)  $\chi_f = \chi_s = 1$ ; whereas  $\chi_c = 100$ . This value for the resonance susceptibility of the cover region was chosen to be consistent with experimental results [65]. No attempt was made to examine possible contributions to an interstitial region between the substrate and film regions.

### 2.5.2 How to Interpret the Numerical Results

For each waveguide (i.e., choice of thickness and film index) and mode combination the direct and cross ratios are calculated via (2.34) through (2.38). The other parameters were considered fixed; they are given in the preceding section. The numerical results are presented in a series of files as described in the beginning of Section 2.5. Before describing the graphical representations of these data, a closer examination of the DR and CR is in order. The DR is the easier-to-interpret indicator of the surface-specificity; it is closely related to the surface-to-total intensity ratio. However, the CR is a better indicator

of the signal-to-noise ratio. The WISOM experimenter usually couples beams to a single waveguide and scans the frequency of one beam looking for a surface resonance by examining the total generated intensity. The top portion of Figure 2.11 shows a representation of such an experiment. Plotted is the total generated field versus the frequency of the scanned beam. Six surface resonances are indicated. In this example they are superimposed on a bulk field of +5 units. For resonances A through C the surface and bulk fields are parallel, whereas in resonances D through F the fields are anti-parallel. Notice that the surface-generated field in resonance B is of the same magnitude as the bulk generated field so that the total field is twice that of the 'background' field. In resonance E, the magnitudes of the two fields are equal, but are anti-parallel giving a zero total field. At the bottom of Figure 2.11, the total field versus increasing surface resonant fields is plotted. The locations where the total field equals the total field of the hypothetical resonances as discussed above are labeled to provide connection with the experimental situation. The total field plot in Figure 2.11 can be compared with the DR shown along the bottom of Figure 2.12 for this situation. Note that when the surface and bulk fields are of the same magnitude, regardless of their direction, the DR is 0.5. If the experimenter encounters a resonance at these points they might appear in the total signal as resonances B or E as shown along the top of the figure. When the surface field overpowers the bulk field by orders of magnitude, the DR approaches one. Representing situations where the total intensity is zero, as in hypothetical resonance E, is more challenging. There are no graphical presentations which provide a quick and meaningful interpretation of this circumstance. To ascertain this condition and quantify if the total signal will be larger or smaller when a surface resonance is encountered the CR may be examined. Figure 2.13 shows the CR for the situation in Figure 2.11 where it is plotted against an increasing surface field. When the surface field is of the same magnitude as the bulk, the CR is +2 when they are parallel (resonance B) and -2 when they are anti-parallel (resonance E). Whenever the CR is -2, the total intensity is zero, however the individual fields from the surface and bulk may be large. Working near this point is advantageous to the

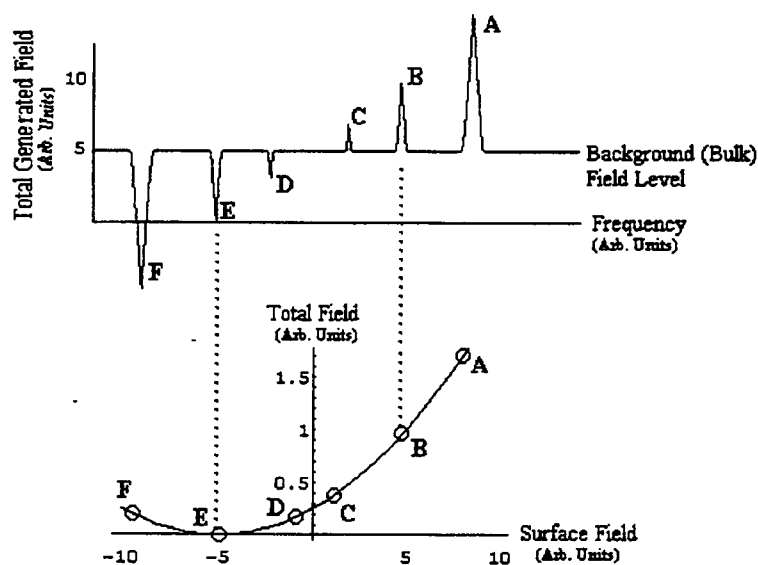


Figure 2.11: Plots of the total field versus frequency for resonances A-F and the total intensity versus surface field.

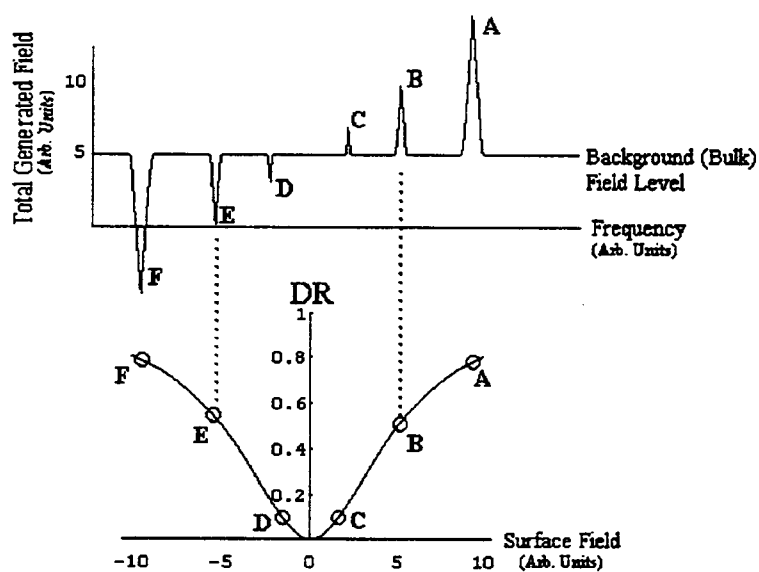


Figure 2.12: Plots of the total field versus frequency for resonances A-F and the DR versus surface field.

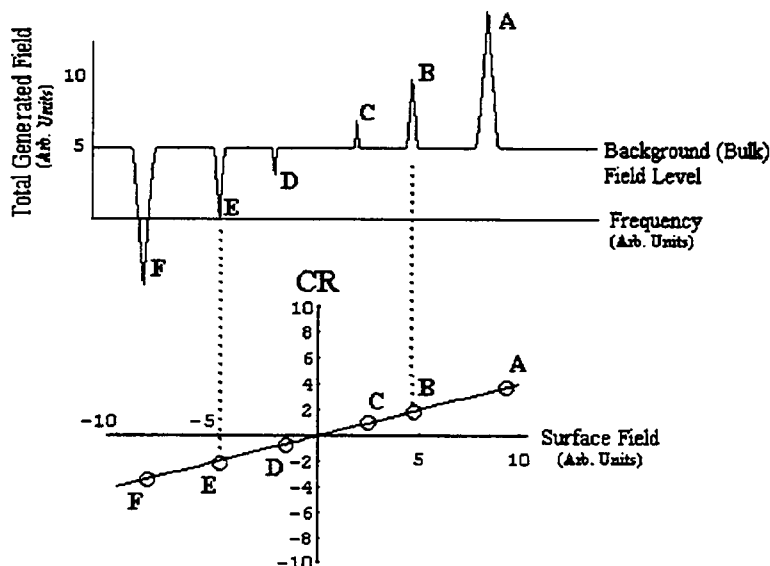


Figure 2.13: Plots of the total field versus frequency for resonances A-F and the CR versus surface field.

experimenter as the signal-to-noise ratio is large. This is different from the case where both fields are very small and the CR is zero. Finally, it is noted that the larger the magnitude of the CR, the better the surface-specificity.

Graphical representations of the DR and CR versus the waveguide parameters were produced for quick comparisons. A series of plots for each mode combination is provided; each plot in the series represents a different difference frequency,  $\Delta$ . These plots are meant to be used as guides to designing waveguides yielding good surface-specificity under a certain mode combination and frequency regime. These are three dimensional plots of film index versus film thickness; the third dimension, the direct or cross ratio, is shown in gray scale. The key to the gray scale is shown along the bottom of each plot. The scale key is marked in the example plot, Figure 2.14. This example clearly labels the film index (along the horizontal) and the film thickness along the vertical axis.

The information bar indicates the ratio type (either direct ratio or cross ratio), the hierarchical data format (hdf) file name used by Spyglass to generate the plot, and the mode combination number. The file name has encoded in it the ratio type (D for DR and C for CR), the difference frequency,  $\Delta$ , which is indicated after the index 'bg'. For instance, the difference frequency for Figure 2.14 is  $1800\text{ cm}^{-1}$ . Each mode combination is given a number; the definitions for the mode number can be found the Mathematica notebook *cars1.ma*. The mode combination for the plot shown in Figure 2.14 is mode number 01, that is,  $[v_1, v_2, v_3, v_4] = [0,0,0,1]$ . Also indicated in Figure 2.14 are three areas of the plot. One area has a low DR; this is an unfavourable ratio. Waveguides and mode combinations yielding good surface-specificity allow for the cancellation of generated signals in the film, leaving the signal only from the surface. When the DR is exactly one, the interference condition is complete. In this case, the observed signal comes only from the cover. If the DR is zero, there is no contribution from the cover region. The region marked as 'unfavourably low', has a DR of less than 0.2. When the cover contribution to the field is half of the total contribution, the DR is 0.25 (i.e., 0.5 squared). When the field contribution from the cover comprises 75% of the total field, the DR is 0.56 (i.e., 0.75 squared). The region labeled 'moderately favourable DR' on Figure 2.14 indicates DRs that represent conditions when the cover contribution is about 75% of the total field. The region of the plot marked 'favourable DR' has a surface-to-total field contribution ratio of greater than 95%, that is the DR is 0.9 or larger. Figure 2.14 indicates that a waveguide with a film index of 1.8 and a thickness of 550 to 600 nm will have a very favourable cover-to-film signal ratio for the  $[0,0,0,1]$  mode with a difference frequency equivalent to  $1800\text{ cm}^{-1}$ . Another waveguide with an index of 1.75 and thickness 650 nm using the same mode and difference frequency will also produce surface-specific results. Note that a plot of the cover-to-film ratio would actually be harder to interpret since this ratio approaches infinity when the cover contribution approaches 100%. The simple bounds of the DR is ample justification to use it as the interpretive tool over the cover-to-film ratio.



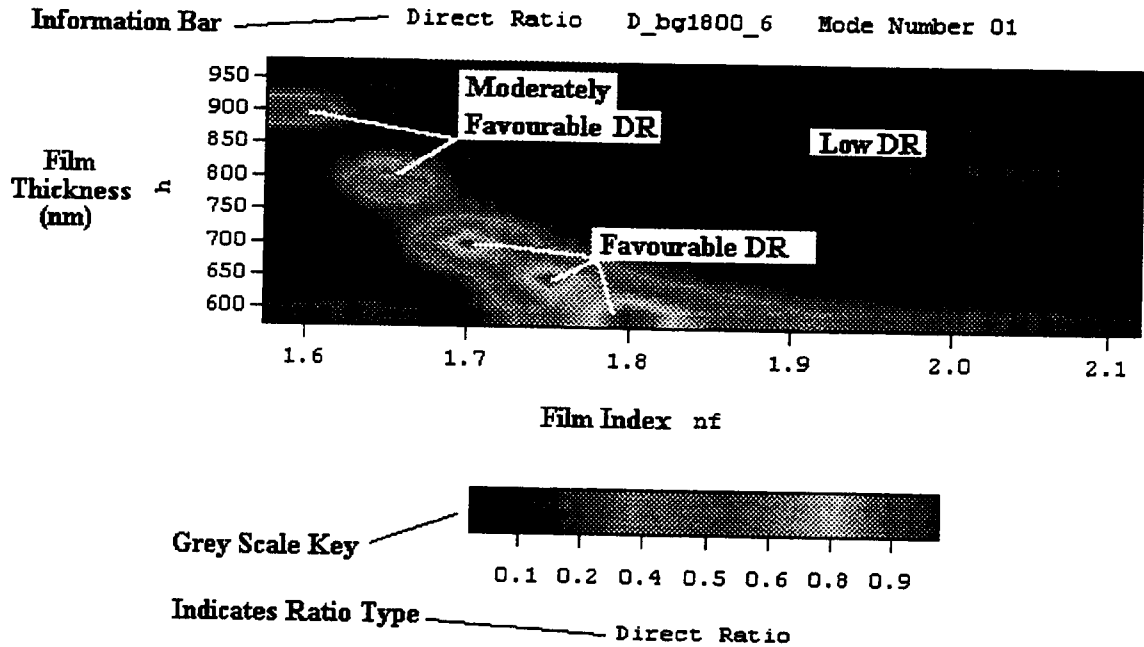


Figure 2.14: An annotated example plot of the direct ratio.

An example of a cross ratio plot is shown in Figure 2.15. The information bar indicates these results are for a waveguide experiment using  $\Delta = 1400 \text{ cm}^{-1}$  and mode combination 13, which is the combination [1,1,1,2]. In the plot, represented by white regions, are situations where the magnitude of the CR is less than one. See the region labeled 'Low CR' in Figure 2.15. These are regions for which the cover-to-bulk field ratio is under 0.5. Regions where cover-to-bulk field ratio is greater than 0.5 are seen as darker regions, darker gray indicating the greater ratio. Roughly half of the gray regions in Figure 2.15 indicate that the phase of the film contribution to the field, relative to the cover phase, is zero. Since the total signal is larger than the bulk signal alone, this condition has been labeled as the 'Favourable CR, Larger Total Field' region in Figure 2.15. In the colour renditions of the CR plots enclosed on the CD-ROM, these regions are tagged in red. When the relative phase is 180 degrees, the total field is less than the bulk alone. This situation has been indicated in Figure 2.15 as a region of 'Favourable

CR, Smaller Total Field'; it is represented in blue on the colour plots. In many cases including this example, a clear white line separates regions of zero and 180 degree phase difference. Normally, the CR is large (greater than 20) when the DR is near one. Hence, the DR can be used almost as the sole tool in studying WISOM experiments. When designing an experiment, the DR can be used to optimize the conditions, but then the CR should be checked to verify it is large. Furthermore, the experimenter can predict whether the signal will increase or decrease when a surface resonance is encountered by examining the CR.

Careful examination of the examples of Figure 2.14 and Figure 2.15 reveals some unusual shapes which are artifacts of the graphical analysis. Scalloped edges in both the DR and CR plots are caused by interpolation errors. Rectangular regions of black in the CR plots occur when the CR exceeds the limits of the plot. Likewise rectangular regions in the DR appearing either black or white occur for the same reason.

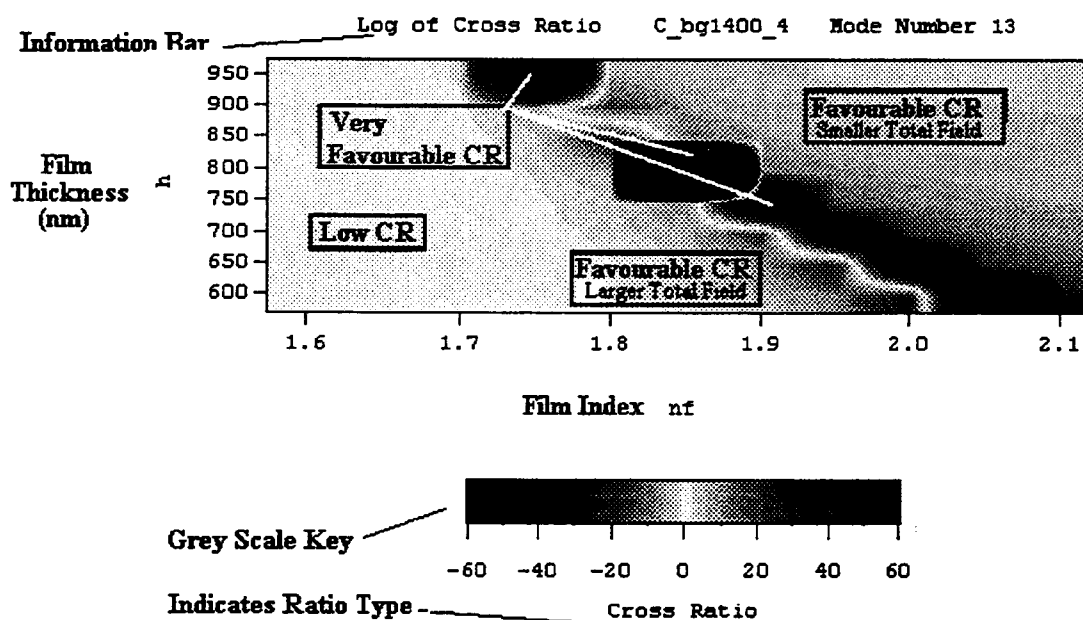


Figure 2.15: An annotated example plot of the cross ratio.

### 2.5.3 Results of the Numerical Analysis

Although all mode combinations that have the potential to yield favourable WISOM conditions were analyzed, only six will be discussed here. These are the six combinations that give the best DR, CR, tuning range, and phase-matching options out of the full eighteen possible combinations. Out of the possible eighteen mode combinations, eight were found to be devoid of the interference condition; these include mode combinations:  $[1,0,0,1]_{\text{MN02}}$ ,  $[2,0,0,1]_{\text{MN03}}$ ,  $[0,1,0,1]_{\text{MN04}}$ ,  $[2,1,0,1]_{\text{MN06}}$ ,  $[0,2,0,1]_{\text{MN07}}$ ,  $[1,2,0,1]_{\text{MN08}}$ ,  $[2,2,1,1]_{\text{MN16}}$ , and  $[2,1,2,1]_{\text{MN18}}$ . The software uses a code, or mode combination number (MN), to designate each mode combination. In the notation above, the mode combination designations appear in the square brackets. The corresponding mode combination number used in the Mathematica notebooks appears as a subscript. Rather than use the four digit mode designation, the mode number is used by the software to identify mode combination and the data files associated with that mode combination. Other combinations less impressive than the six best combinations include:  $[2,0,1,1]_{\text{MN11}}$ ,  $[1,1,1,0]_{\text{MN12}}$ ,  $[2,1,1,1]_{\text{MN14}}$ ,  $[1,2,1,1]_{\text{MN15}}$ . These mode combinations are not reviewed here. The results are available on the enclosed CD-ROM. Each mode combination has its own sub-directory in the *wscars* directory. The raw data is available in text and as a Microsoft Excel (V. 3.0) file; the DR and CR are available both as *hdf* files and in bitmap (BMP) form each according to difference frequency index. For example, mode combination  $[1,1,1,0]_{\text{MN12}}$ , is located in the sub-directory *cars0m12*. The raw data is contained in a single file available in both formats. The plots are identified by a frequency index. For instance, for MN12 and difference frequency  $1000 \text{ cm}^{-1}$  (frequency index 02) have the identifiers *d0m1202* for the DR and *c0m1202*.

The DR and CR plots for the six mode combinations reported are shown in Figure 2.16 through Figure 2.48. Note that each figure has the DR and the CR combined on each figure. These six promising mode combinations are:  $[2,2,0,1]_{\text{MN09}}$ ,  $[1,0,1,1]_{\text{MN10}}$ ,  $[1,1,1,2]_{\text{MN13}}$ ,  $[0,0,0,1]_{\text{MN01}}$ ,  $[2,0,2,1]_{\text{MN17}}$ , and  $[1,1,0,1]_{\text{MN05}}$ . Note that two of these

combinations are four-beam mode combinations. The next several paragraphs discuss each mode combination in further detail.

The DRs and CRs for mode combination [2,2,0,1] is shown in Figure 2.16 through Figure 2.21. Each figure contains the results for the difference frequency,  $\Delta$ ,  $400 \text{ cm}^{-1}$  larger than the previous figure. Figure 2.16 shows the results for the difference frequency  $1000 \text{ cm}^{-1}$ . For film indices above 1.85, the DR and CR are large (between 0.999 and one) and are relatively insensitive to the difference frequency. This is especially true of film thicknesses between 750 and 850 nm. Hence, mode combination [2,2,0,1] has an amazingly large tuning range, that is, the range over the difference frequency for which interference condition is quite favourable. For instance, a waveguide with index 1.95 and thickness 850 nm has a DR of 0.9999 (four nines) or greater and a CR of more than 30 in magnitude. A clear demarcation can be seen between waveguide parameters that give parallel or anti-parallel surface and bulk fields. To the left of this white boundary the surface and bulk fields add constructively. There are very remarkable cover-to-film field ratios for certain frequencies with specific waveguides and using the [2,2,0,1] mode combination. For a waveguide with a film index of 1.95 and thickness 850 nm, using a difference frequency of  $1000 \text{ cm}^{-1}$ , for example, the DR is seven nines and the CR is  $3.9 \times 10^3$ . Other waveguides give similarly effective interference conditions for a number of difference frequencies: for  $\Delta = 1200 \text{ cm}^{-1}$ , an index of 1.9 and thickness 800 nm gives a DR of six nines and a CR of  $-2.3 \times 10^3$ ; for  $\Delta = 1800 \text{ cm}^{-1}$ , an index of 1.9 and thickness 700 nm gives a DR of eight nines and a CR of  $-4.1 \times 10^3$ ; and for  $\Delta = 2800 \text{ cm}^{-1}$ , an index of 2.1 and thickness 600 nm gives a DR of five nines and a CR of 969. Notice, that for any frequency, a waveguide with nominal thickness of 800 nm and index 1.95 has a DR above 0.99 and a CR of at least 40. However, for this mode combination almost any waveguide with indices smaller than 1.7 is ineffective for WISOM experiments. As will be reviewed in the following section, this mode combination also offers many opportunities to phase-match the FWM process. Moreover, many of these phase-matching conditions have larger interaction regions. That is, the overlap is large since the crossing angles for

which phase-matching occurs are small. As will be seen later, most phase-matching conditions require that the azimuthal (i.e., crossing) angles be 90 degrees apart (the beams cross perpendicularly). This makes for small interaction regions.

Mode combination [1,0,1,1] holds promise for WISOM experiments for smaller difference frequencies. Figure 2.22 through Figure 2.27 show the DRs and CRs for difference frequencies from  $800\text{ cm}^{-1}$  to  $1800\text{ cm}^{-1}$  in  $200\text{ cm}^{-1}$  steps. Unfortunately, the interference condition is rather sensitive to frequency. A waveguide with film index 1.85 and thickness 750 nm has a DR of five nines and a CR of 652 for  $\Delta = 800\text{ cm}^{-1}$ . The same waveguide at  $\Delta = 1000\text{ cm}^{-1}$  has a DR of 0.957 and a CR of -9.3. For the set of frequencies chosen, and difference frequency  $\Delta = 1200\text{ cm}^{-1}$ , there are two possible waveguides that give remarkable surface-specificity. These are waveguides with indices of 1.70. For the chosen input frequencies, such a waveguide with film thickness 700 nm can have a DR of ten nines and a CR of  $+9.3 \times 10^3$ . Another similar waveguide having thickness 600 nm can have a DR of nine nines and a CR of  $-3.3 \times 10^3$ . This same waveguide has a DR of larger than 0.95 and a respectable CR of -8.6 at  $\Delta = 1400\text{ cm}^{-1}$ . For larger difference frequencies, those above  $1800\text{ cm}^{-1}$ , the effectiveness of this mode combination for WISOM work is lost. Again, the left side of the CR plots show that the surface and bulk fields are parallel, however for  $\Delta$  greater than  $1400\text{ cm}^{-1}$  all waveguide parameters give only anti-parallel fields. This mode combination also allows for many phase-matching options that have larger interaction regions.

Figure 2.28 through Figure 2.33 show results for mode combination [1,1,1,2]. This is considered a favourite of the three-beam mode combinations used by experimental researchers [27], [26], [30], [31]. These plots start with the results for  $\Delta = 1000\text{ cm}^{-1}$  and increasingly larger difference frequencies in  $400\text{ cm}^{-1}$  steps. This mode combination is far more sensitive to the difference frequency than the combinations previously discussed. For instance, a waveguide with a film index of 1.9 and thickness of 600 nm at  $\Delta = 1800\text{ cm}^{-1}$  has a DR of 0.997 and a CR of +36. However, at  $\Delta = 1400\text{ cm}^{-1}$  or  $2200\text{ cm}^{-1}$ , this

waveguide has a much smaller DR (0.86 and 0.94, respectively) and CR (+5 and -8, respectively). For  $\Delta = 1400 \text{ cm}^{-1}$ , a better waveguide to use would be one with thickness 800 nm and index 1.85; this gives a DR of five nines and a CR of 829. Likewise, if the experimenter wishes to tune to difference frequency of  $\Delta = 2200 \text{ cm}^{-1}$ , a different waveguide should be used than that used at either  $\Delta = 1400 \text{ cm}^{-1}$  or  $\Delta = 1800 \text{ cm}^{-1}$ . The best waveguide for this higher frequency is one with a film index of 1.75 and thickness 750 nm; it yields a DR of five nines and a CR of -269. Hence, in using this mode combination, the experimenter would be required to change waveguides while tuning over this  $800 \text{ cm}^{-1}$  range. This is certainly undesirable since many other factors between waveguides may be different other than the film index and thickness. Such factors include the surface roughness and the mode coupling efficiency.

Next in the series of mode combinations discussed here is the combination [0,0,0,1]. This combination is another of the three-beam mode combinations commonly used by experimentalists. Although it does not provide as great a surface-specificity, it is an easy combination of modes to couple into. The efficiency of coupling is related to the mode; lower numbered modes are coupled into with greater efficiency. Therefore, this mode combination is quite attractive. Surface-specificity is sacrificed for signal strength. From Figure 2.34 through Figure 2.39, it can be seen that the combination is very sensitive to the difference frequency. The plots show the difference frequency from  $\Delta = 800 \text{ cm}^{-1}$  to  $\Delta = 1800 \text{ cm}^{-1}$  in  $200 \text{ cm}^{-1}$  steps. In that range of difference frequencies, there is no single waveguide that would yield good WISOM conditions. Again, the experimenter would be required to change out waveguides during any tuning experiment. For instance, at  $\Delta = 800 \text{ cm}^{-1}$  surface-specificity will be best using a waveguide that has an index of 2.05 and thickness of 750 nm (the DR is two nines and the CR is -25). When tuning up to  $\Delta = 1000 \text{ cm}^{-1}$ , the best waveguide would be one with an index of 2.05 and an thickness of 650 nm. This is the best that this mode combination is capable of. This is because small number modes like modes 0 and 1 have small amplitudes at the surface. This waveguide and frequency gives a DR and a CR of 0.979 and +13.5, respectively. Above  $\Delta = 1800$

$\text{cm}^{-1}$ , the mode combination loses its ability to provide good surface-specificity regardless of the choice of waveguides.

The fifth mode combination reviewed here is [2,0,2,1]. A single waveguide of index 2.1 and thickness 600 nm gives reasonable surface-specificity over a modest tuning range of  $1800 \text{ cm}^{-1}$ . This is shown in Figure 2.40 through Figure 2.45 where the DRs and CRs are shown from  $800 \text{ cm}^{-1}$  to  $1800 \text{ cm}^{-1}$  in steps of  $200 \text{ cm}^{-1}$ . At  $\Delta = 800 \text{ cm}^{-1}$ , the DR is 0.948 and the CR is -8.5. As the difference frequency climbs, the specificity drops slowly; at  $\Delta = 1400 \text{ cm}^{-1}$ , the DR and CR are 0.885 and -5.5, respectively. As can be seen from these figures, this waveguide is the best for this mode combination for any tuning value. No matter what the waveguide parameters, this mode combination produces anti-parallel surface and bulk fields. It is interesting to note that for the lower difference frequencies, the cover-to-film field ratio is rather insensitive to the waveguide parameters. Consequently for these values of  $\Delta$ , almost any waveguide gives reasonable results. This mode combination is also very sensitive to the azimuthal angles in regards to phase-matching with respect to changes in the frequency. As will be seen in the next section, if the cover and film have different dispersions, this sensitivity can be of great advantage in increasing the surface-specificity.

Mode combination [1,1,0,1] is not noteworthy for its large cover-to-total signal ratios. This mode combination produces only parallel fields. As seen from Figure 2.46 through Figure 2.48, the best waveguides have large indices and thin films. For instance, a waveguide using this mode has a DR of 0.936 and a CR of +7.6 for  $\Delta = 800 \text{ cm}^{-1}$  when it has a thickness of 600 nm and an index of 2.1. Increasing the frequency to  $\Delta = 1400 \text{ cm}^{-1}$ , this waveguide yields a DR of 0.91 and CR of +6.2. What is exceptional about this mode is the probability that phase-matching can be achieved with small azimuthal angles. The experimenter, if favoured with a large dispersion of the film with respect to the cover, may find that WISOM can be achieved with lower DRs and CRs while utilizing this unique phase-matching condition.

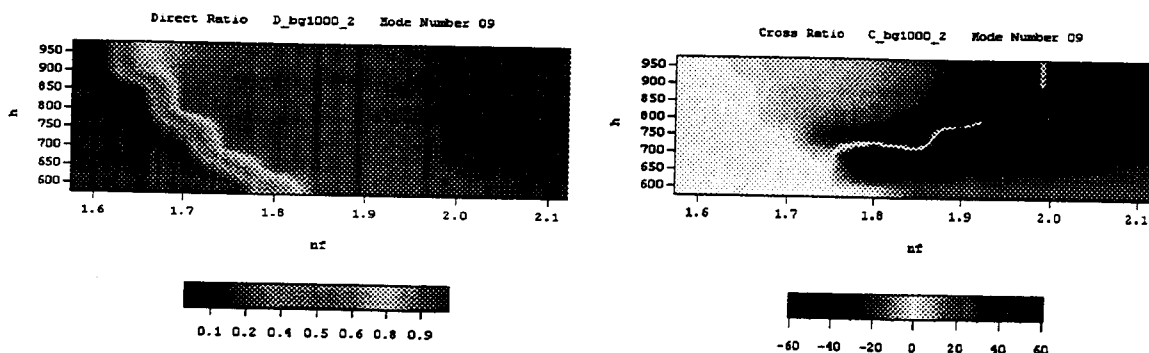


Figure 2.16: The DR and CR for mode combination [2,2,0,1] at  $\Delta = 1000 \text{ cm}^{-1}$ .

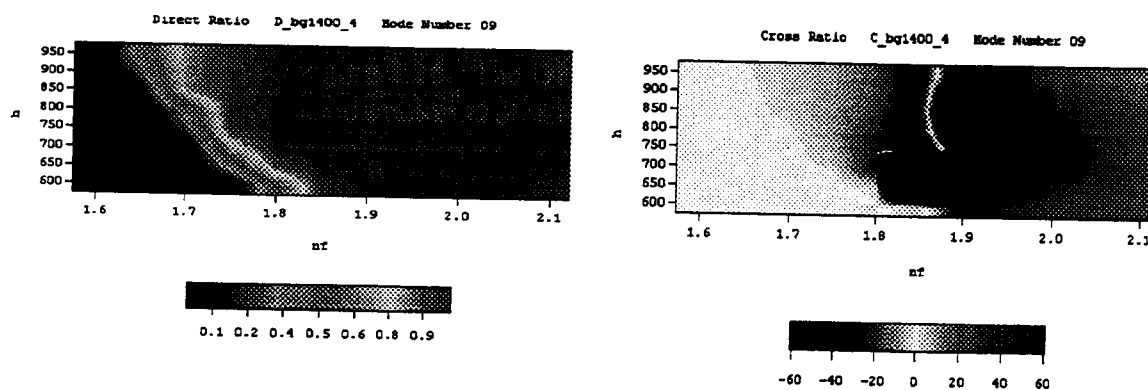


Figure 2.17: The DR and CR for mode combination [2,2,0,1] at  $\Delta = 1400 \text{ cm}^{-1}$ .

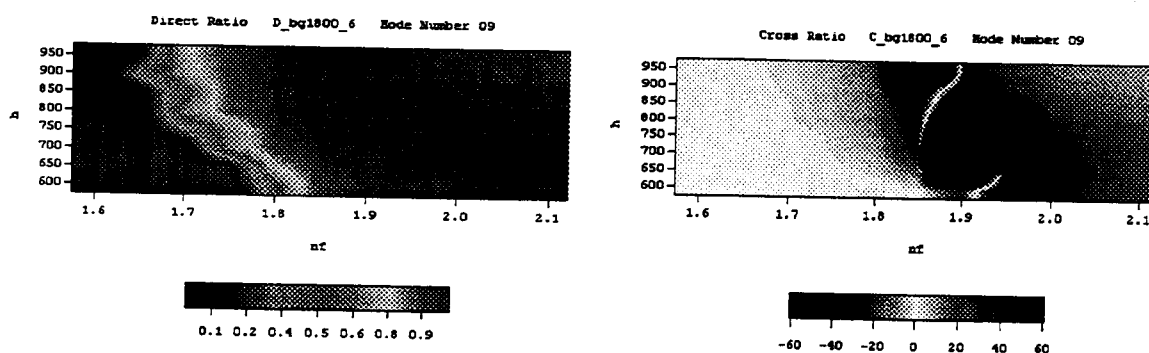


Figure 2.18: The DR and CR for mode combination [2,2,0,1] at  $\Delta = 1800 \text{ cm}^{-1}$ .



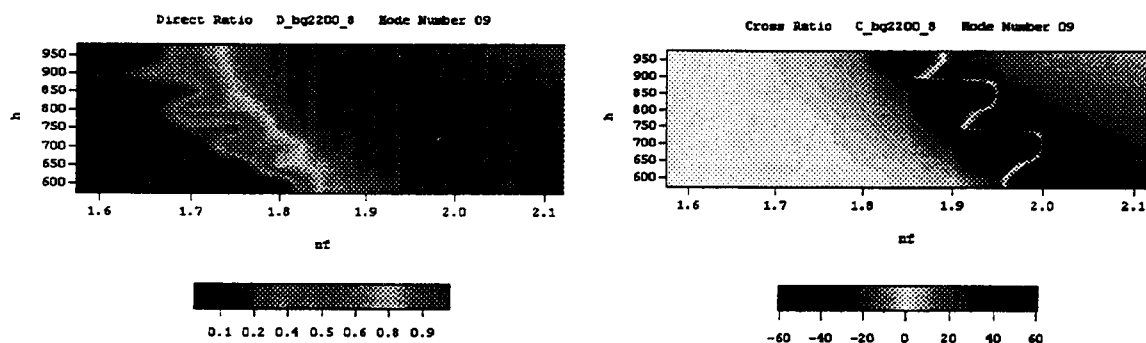


Figure 2.19: The DR and CR for mode combination  $[2,2,0,1]$  at  $\Delta = 2200 \text{ cm}^{-1}$ .

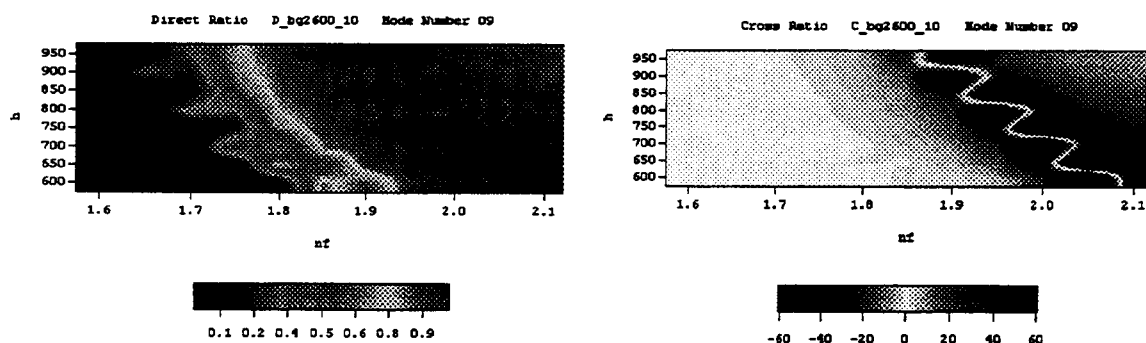


Figure 2.20: The DR and CR for mode combination  $[2,2,0,1]$  at  $\Delta = 2600 \text{ cm}^{-1}$ .

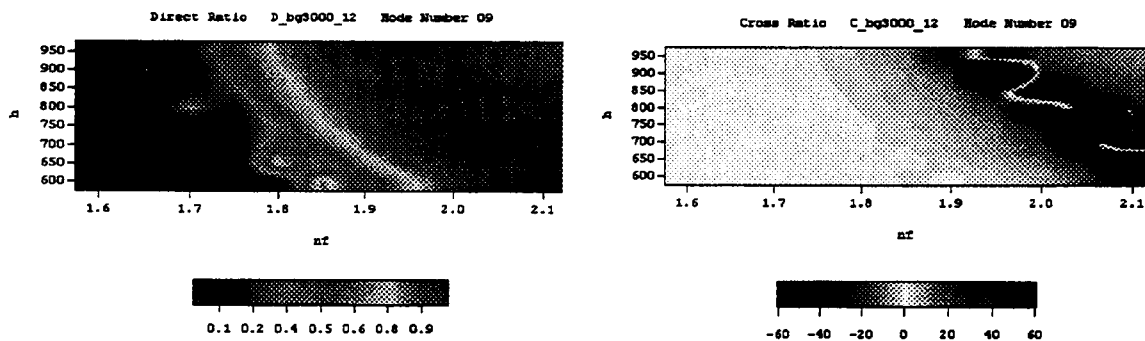


Figure 2.21: The DR and CR for mode combination  $[2,2,0,1]$  at  $\Delta = 3000 \text{ cm}^{-1}$ .

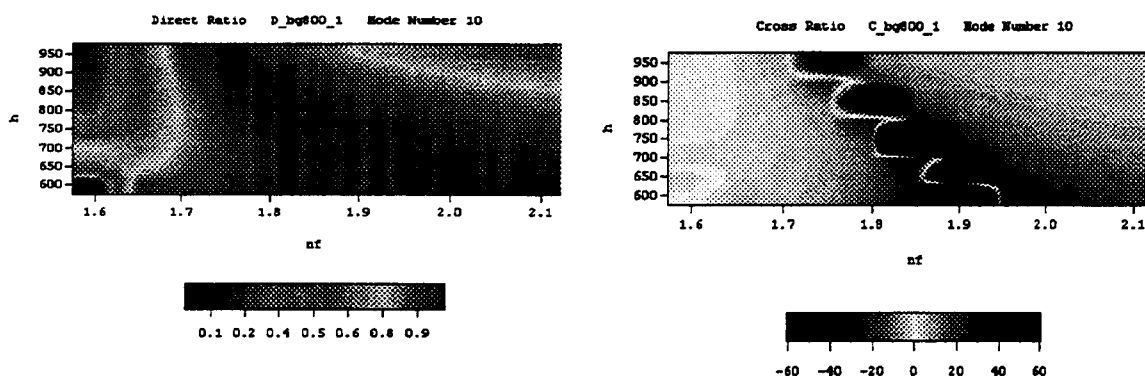


Figure 2.22: The DR and CR for mode combination  $[1,0,1,1]$  at  $\Delta = 800 \text{ cm}^{-1}$ .

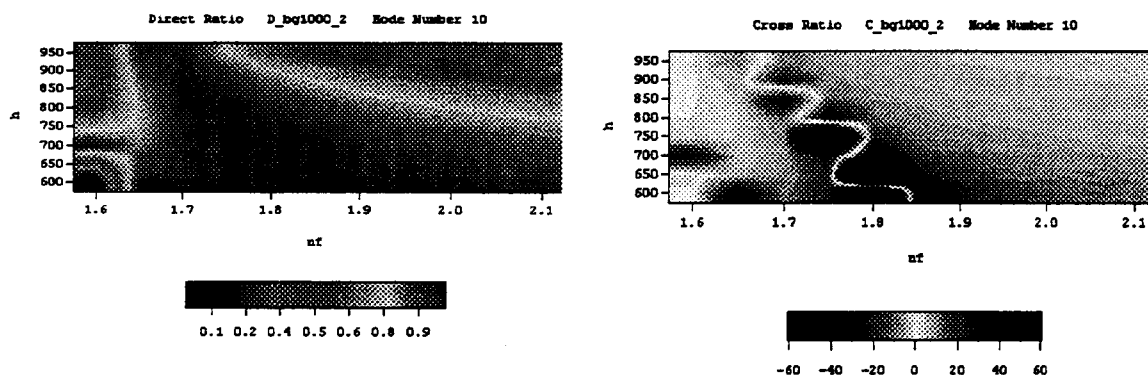


Figure 2.23: The DR and CR for mode combination  $[1,0,1,1]$  at  $\Delta = 1000 \text{ cm}^{-1}$ .

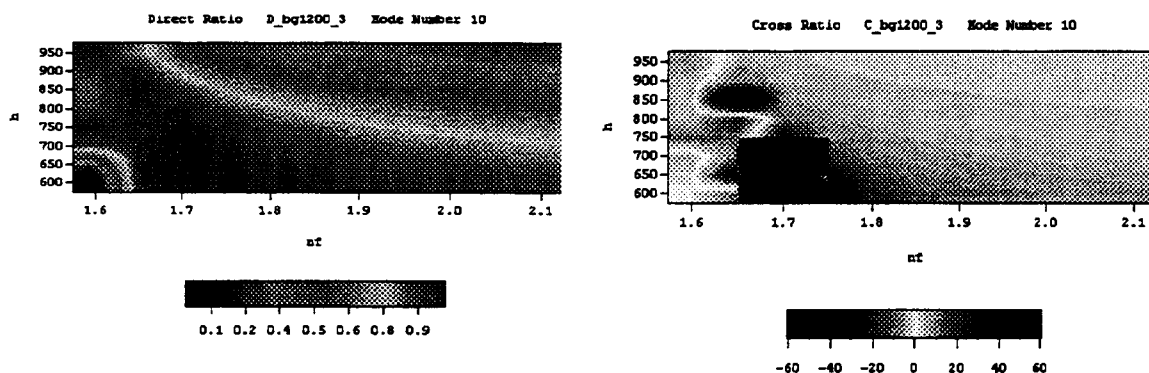


Figure 2.24: The DR and CR for mode combination  $[1,0,1,1]$  at  $\Delta = 1200 \text{ cm}^{-1}$ .

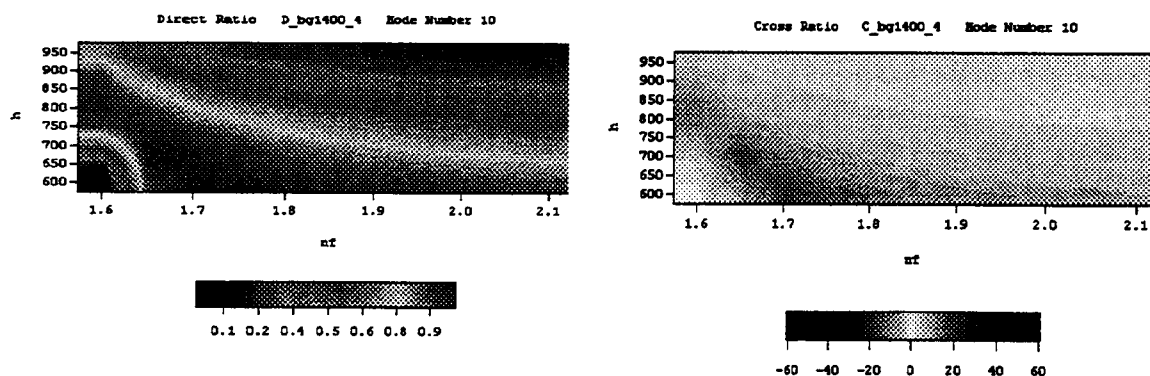


Figure 2.25: The DR and CR for mode combination  $[1,0,1,1]$  at  $\Delta = 1400 \text{ cm}^{-1}$ .

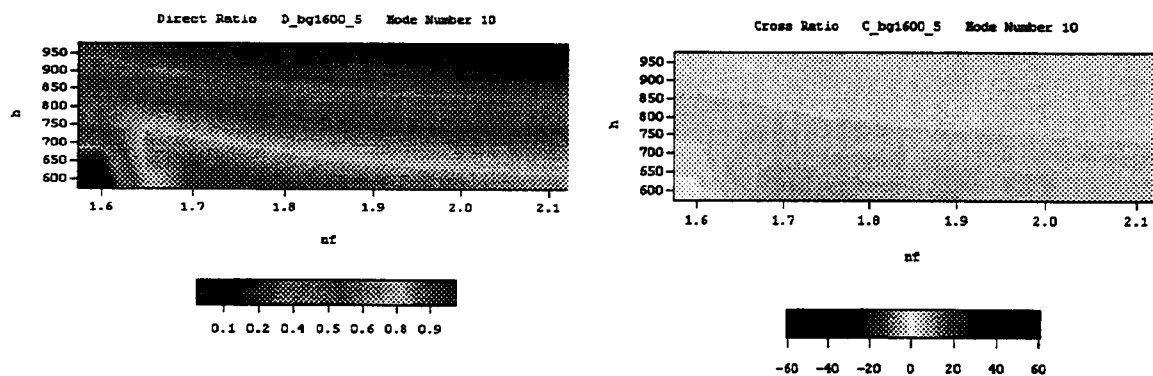


Figure 2.26: The DR and CR for mode combination  $[1,0,1,1]$  at  $\Delta = 1600 \text{ cm}^{-1}$ .

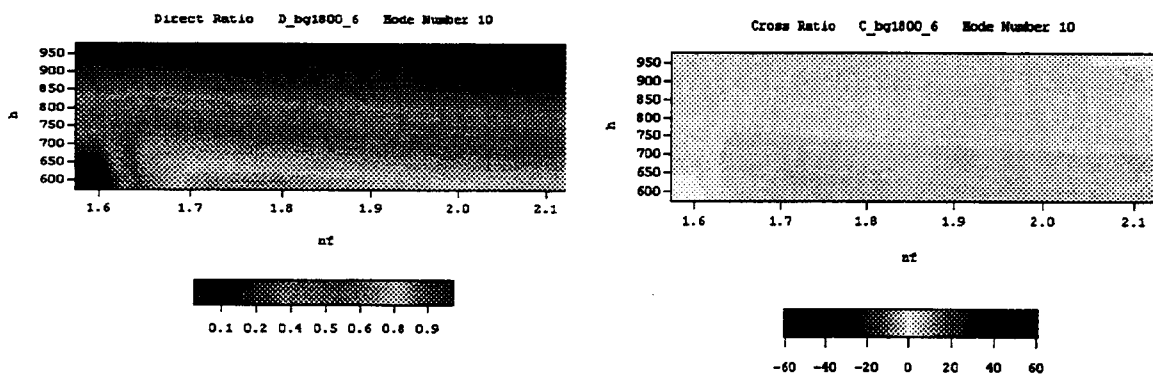


Figure 2.27: The DR and CR for mode combination  $[1,0,1,1]$  at  $\Delta = 1800 \text{ cm}^{-1}$ .

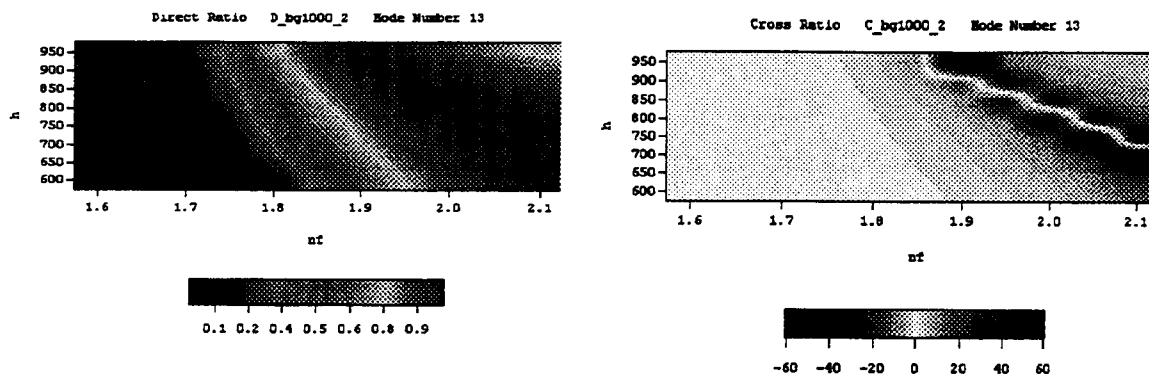


Figure 2.28: The DR and CR for mode combination [1,1,1,2] at  $\Delta = 1000 \text{ cm}^{-1}$ .

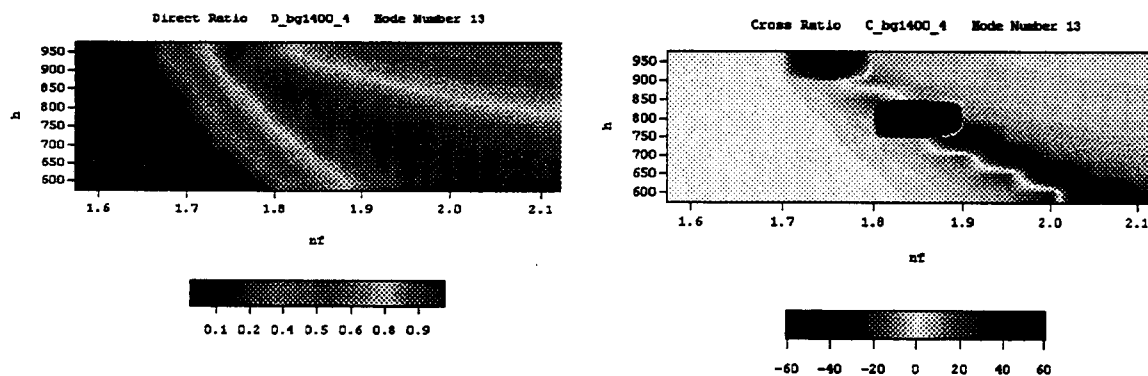


Figure 2.29: The DR and CR for mode combination [1,1,1,2] at  $\Delta = 1400 \text{ cm}^{-1}$ .

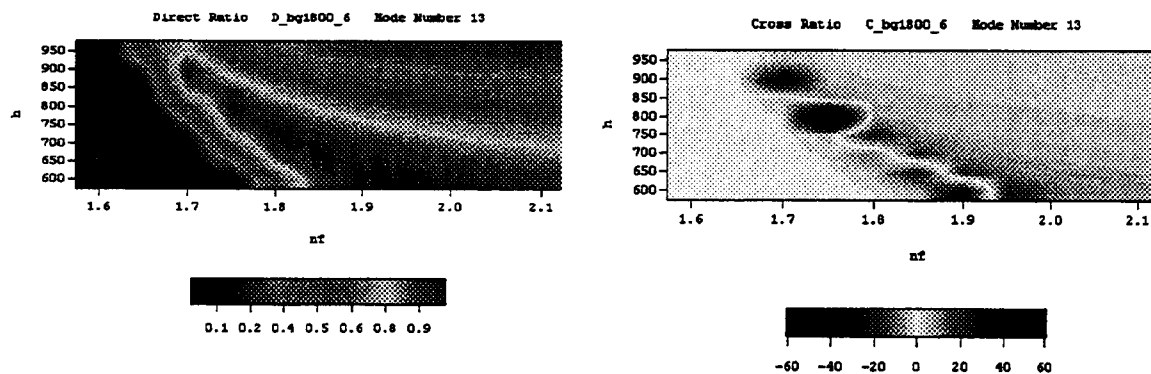


Figure 2.30: The DR and CR for mode combination [1,1,1,2] at  $\Delta = 1800 \text{ cm}^{-1}$ .

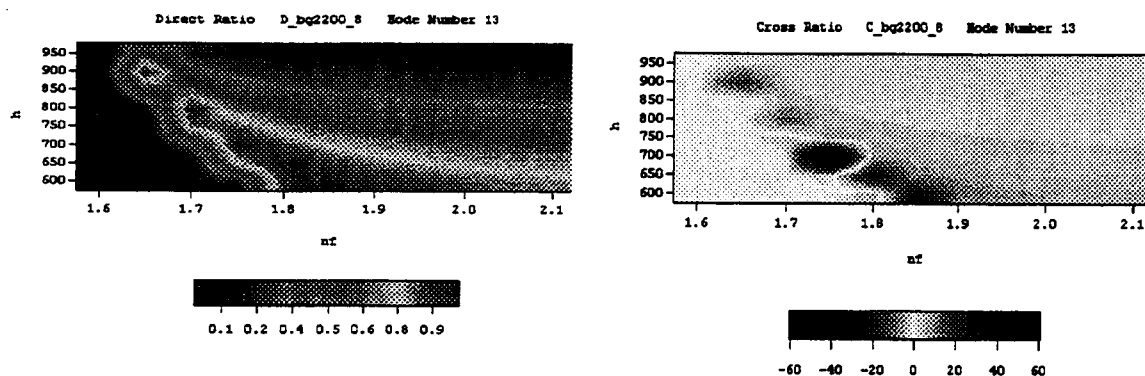


Figure 2.31: The DR and CR for mode combination  $[1,1,1,2]$  at  $\Delta = 2200 \text{ cm}^{-1}$ .

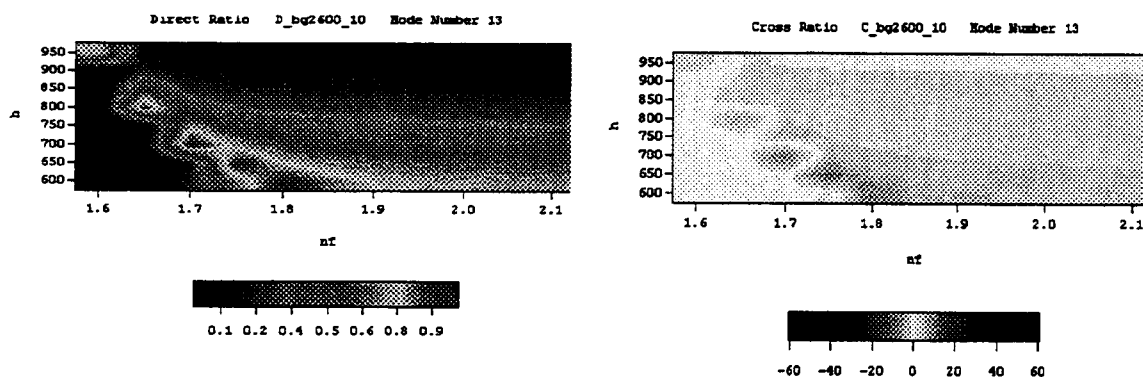


Figure 2.32: The DR and CR for mode combination  $[1,1,1,2]$  at  $\Delta = 2600 \text{ cm}^{-1}$ .

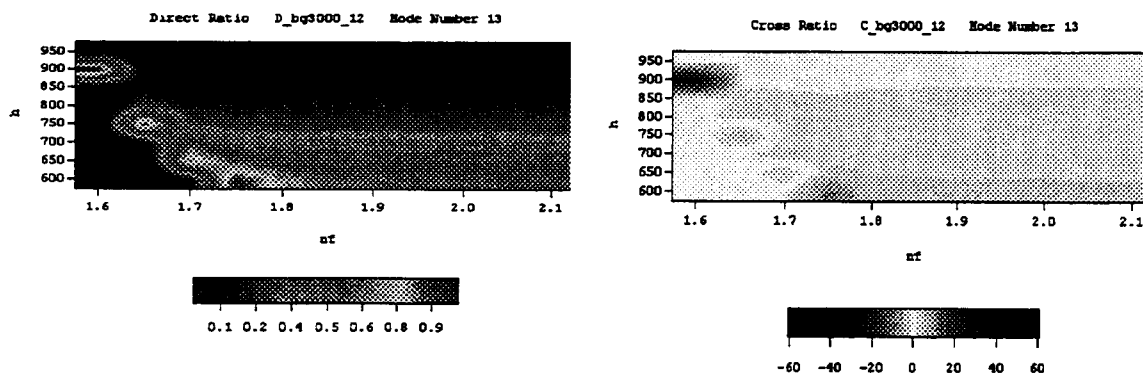


Figure 2.33: The DR and CR for mode combination  $[1,1,1,2]$  at  $\Delta = 3000 \text{ cm}^{-1}$ .

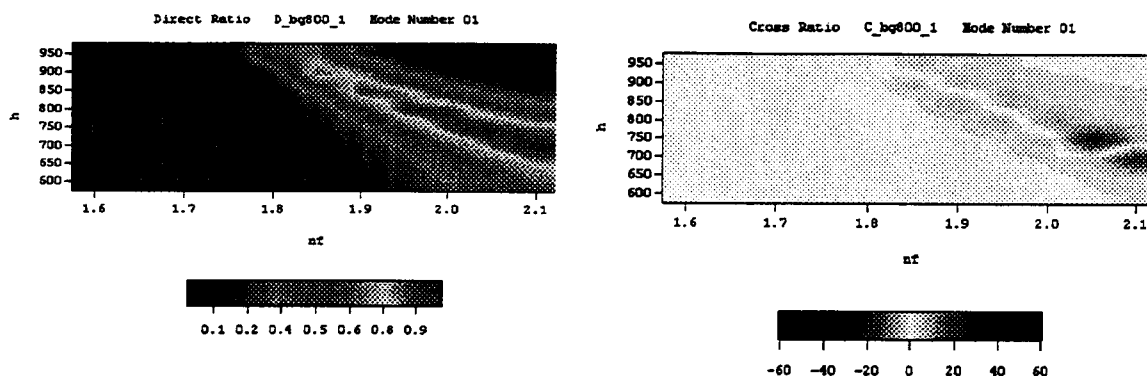


Figure 2.34: The DR and CR for mode combination [0,0,0,1] at  $\Delta = 800 \text{ cm}^{-1}$ .

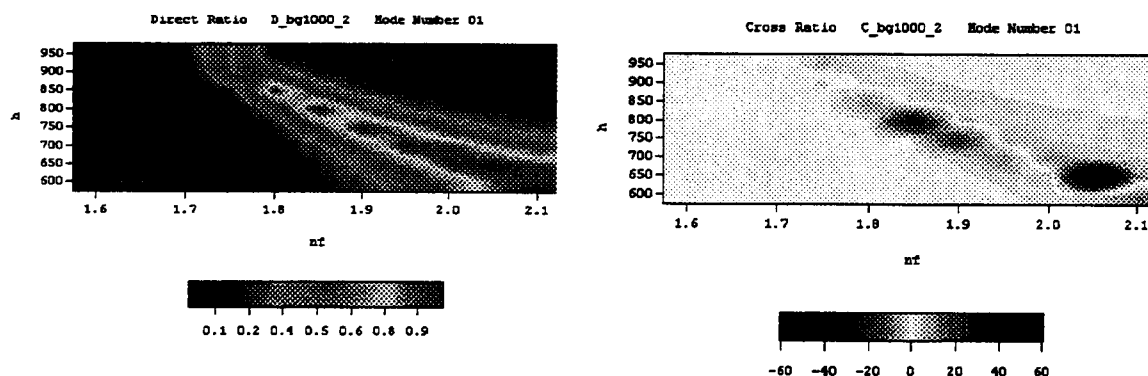


Figure 2.35: The DR and CR for mode combination [0,0,0,1] at  $\Delta = 1000 \text{ cm}^{-1}$ .

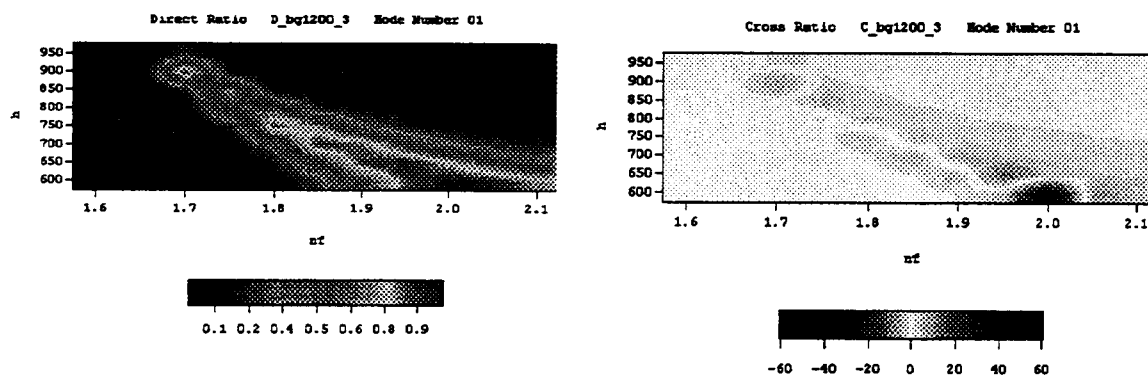


Figure 2.36: The DR and CR for mode combination [0,0,0,1] at  $\Delta = 1200 \text{ cm}^{-1}$ .

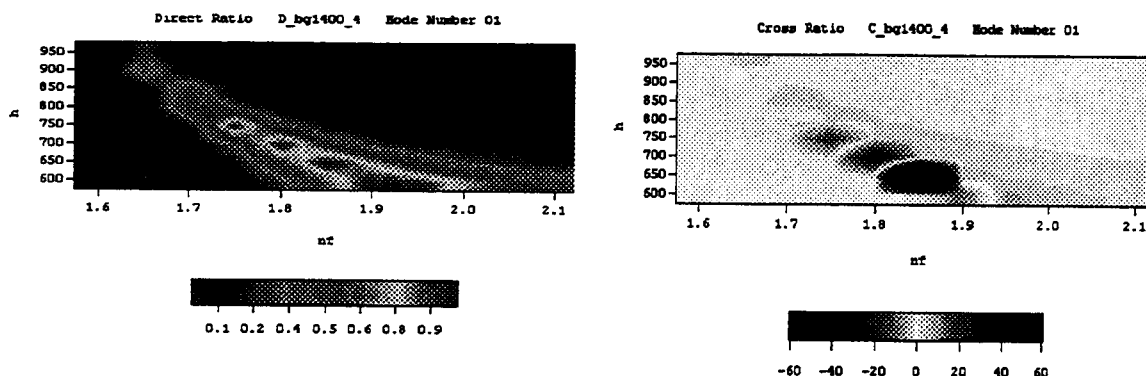


Figure 2.37: The DR and CR for mode combination [0,0,0,1] at  $\Delta = 1400 \text{ cm}^{-1}$ .

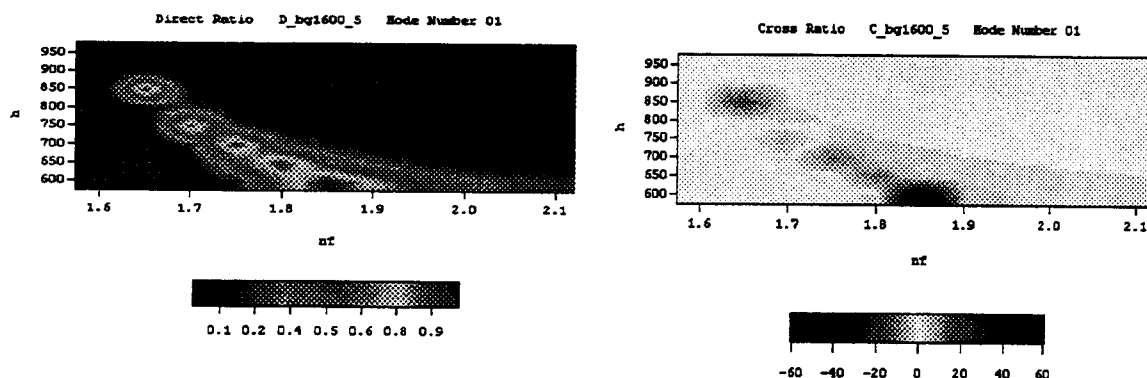


Figure 2.38: The DR and CR for mode combination [0,0,0,1] at  $\Delta = 1600 \text{ cm}^{-1}$ .

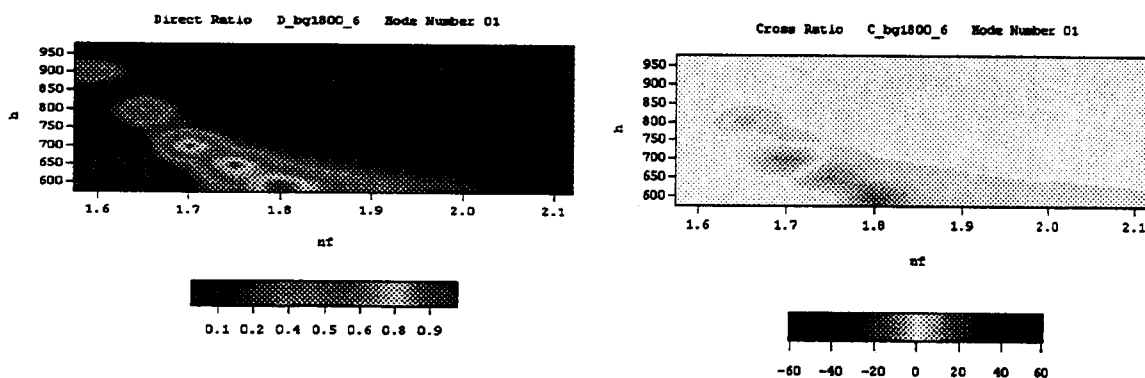


Figure 2.39: The DR and CR for mode combination [0,0,0,1] at  $\Delta = 1800 \text{ cm}^{-1}$ .

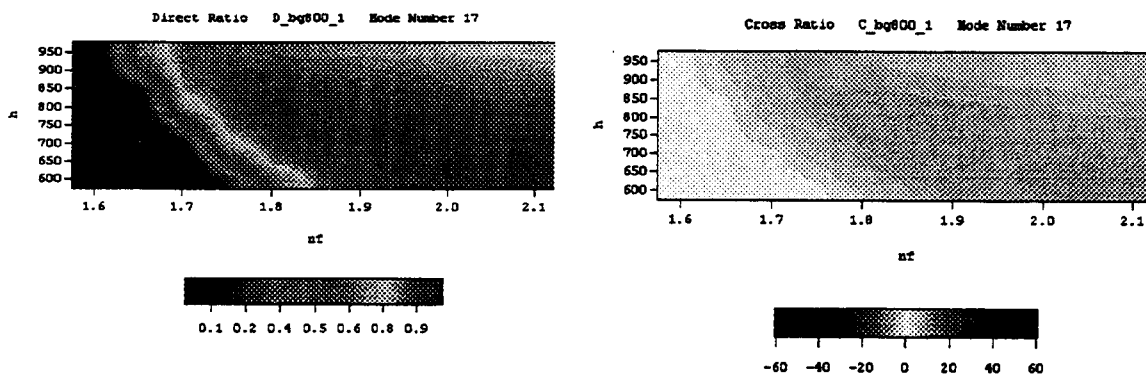


Figure 2.40: The DR and CR for mode combination  $[2,0,2,1]$  at  $\Delta = 800 \text{ cm}^{-1}$ .

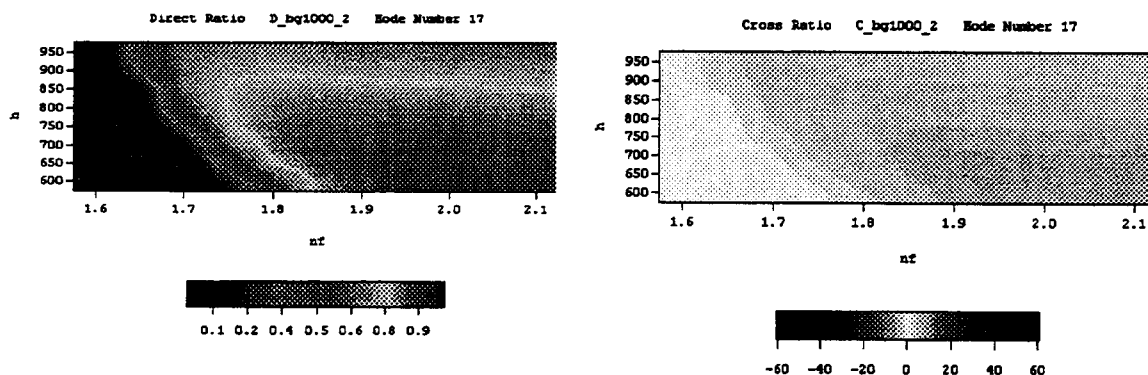


Figure 2.41: The DR and CR for mode combination  $[2,0,2,1]$  at  $\Delta = 1000 \text{ cm}^{-1}$ .

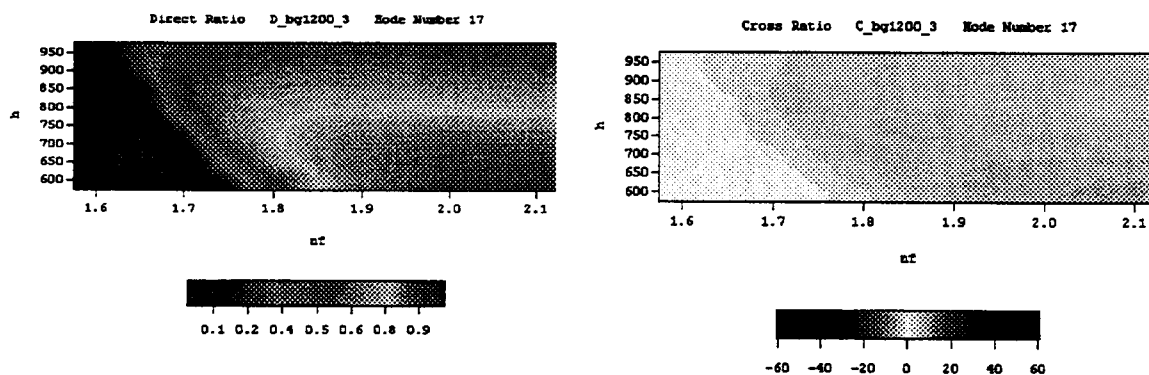


Figure 2.42: The DR and CR for mode combination  $[2,0,2,1]$  at  $\Delta = 1200 \text{ cm}^{-1}$ .



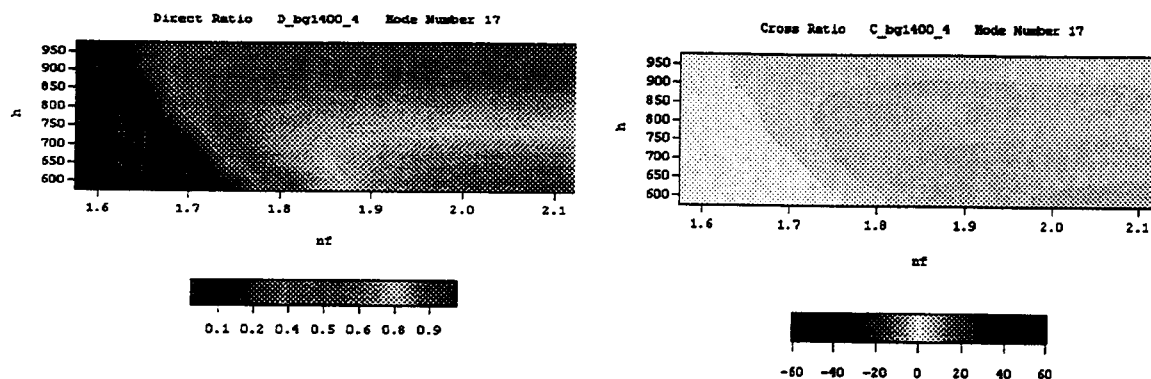


Figure 2.43: The DR and CR for mode combination  $[2,0,2,1]$  at  $\Delta = 1400 \text{ cm}^{-1}$ .

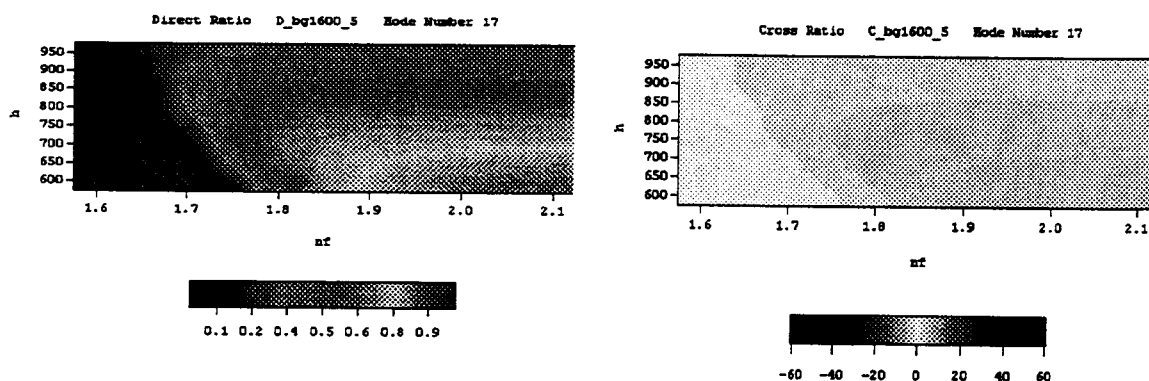


Figure 2.44: The DR and CR for mode combination  $[2,0,2,1]$  at  $\Delta = 1600 \text{ cm}^{-1}$ .

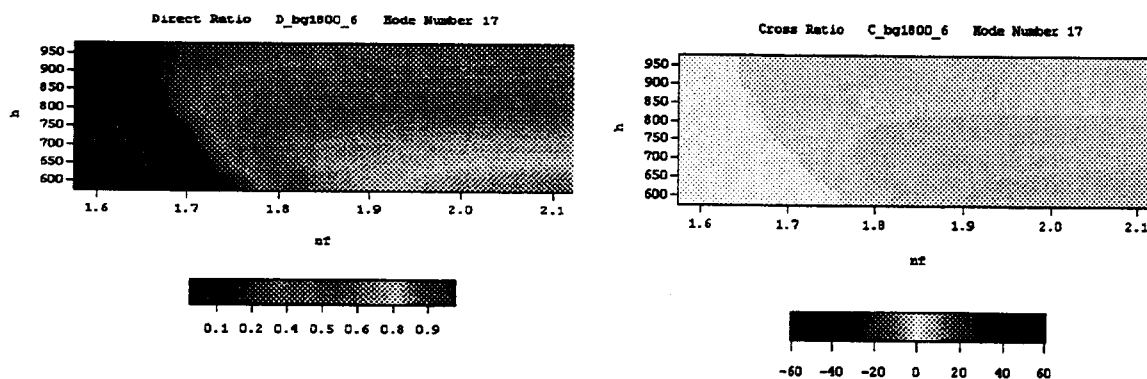


Figure 2.45: The DR and CR for mode combination  $[2,0,2,1]$  at  $\Delta = 1800 \text{ cm}^{-1}$ .

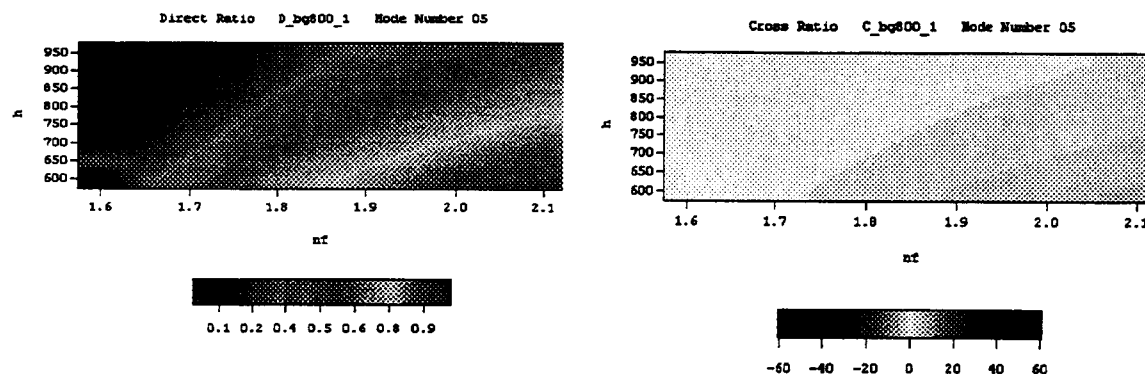


Figure 2.46: The DR and CR for mode combination [1,1,0,1] at  $\Delta = 800 \text{ cm}^{-1}$ .

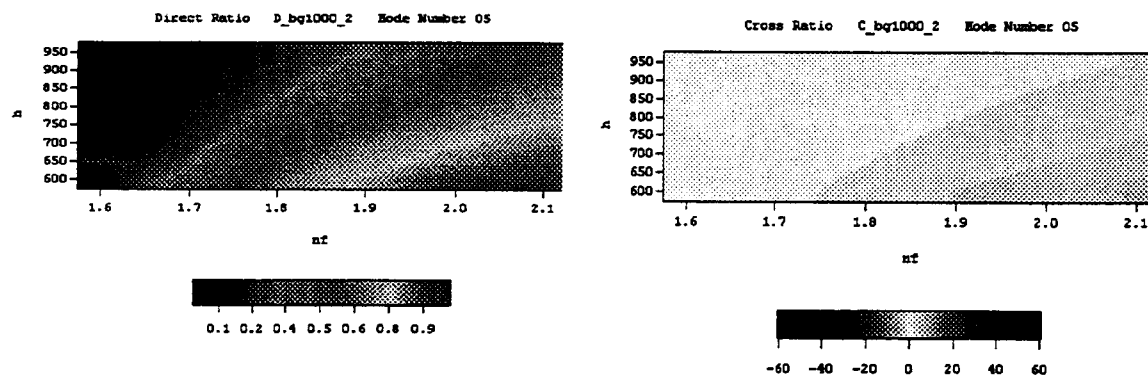


Figure 2.47: The DR and CR for mode combination [1,1,0,1] at  $\Delta = 1000 \text{ cm}^{-1}$ .

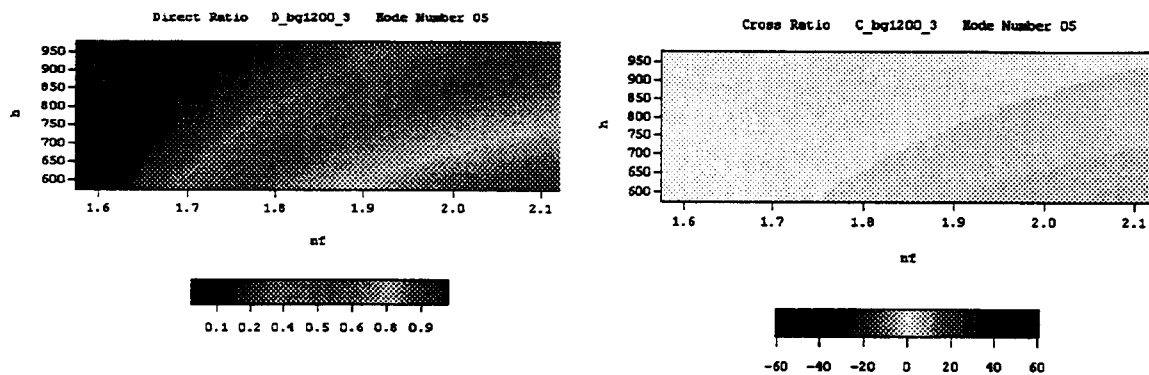


Figure 2.48: The DR and CR for mode combination [1,1,0,1] at  $\Delta = 1200 \text{ cm}^{-1}$ .

## 2.6 Phase-Matching Gaussian Beams in a Waveguide

So far in the analysis of these waveguides, phase-mismatches have been ignored. Recall that the polar angles are determined by the mode combination. Clearly, in order to phase-match the OM process, the azimuthal angles must be chosen. It is likely that phase-matching will not occur when the waves are collinear. Occasionally, as was found from the previous set of calculations, that there is no choice of azimuthal angles that will phase-match the process inside the waveguide. More often, the waves must be nearly perpendicular in order to achieve a zero phase-mismatch. In fact, in over 70% of the choices of waveguides, mode combinations, and difference frequencies; phase-matching occurs when the azimuthal angles are approximately right angles from one another. In 2% of the cases, no phase-matching was possible. In the remaining cases, the azimuthal angles were smaller. Furthermore, in about 18% of the choices, the azimuthal angles were such that the applied waves were under 3 degrees in separation when the mismatch was zero. Experiments where the angular separation of the waves is small are mechanically simplified. Moreover, when the azimuthal angles are small, the interaction volume is large. Hence, these conditions were particularly sought. The exact azimuthal angles for every waveguide, mode combination, and difference frequency are given in the files on the CD-ROM discussed earlier.

In this section, a realistic experiment is considered where three focused Gaussian profile beams are coupled into a waveguide and mix with the proper phase-mismatch according to the coupling (polar) and separation (azimuthal) angles. To find the intensity of the output beam, a volume integral must be done for each wavevector in each beam. Results show that Gaussian applied beams mixed together to form an expanding Gaussian output beam as expected. First, representations of a Gaussian beam and the integration

yielding the proper intensity are discussed. This is followed by an analysis of the intensity as the azimuthal angles are varied from no angular separation to the separation for which there is no phase-mismatch. Realizing a smaller angular separation gives a larger interaction volume, the question to be answered is: What is the larger competing factor to the overall cover intensity, the interaction volume or the phase-mismatch?

### 2.6.1 Deriving the Phase-Mismatch of WISOM Using Focused Gaussian Beams

To best simulate a laboratory experiment, focused Gaussian beams were considered instead of plane waves in the calculations for the phase-mismatch of WISOM experiments. Typically, the source waves for WISOM experiments are lasers with Gaussian profiles. Furthermore, these waves are focused with 0.3 to 1.0 meter focal length lenses prior to coupling to the waveguide (see Figure 2.7 and Figure 2.8). This increases the field amplitude at the point of crossing (the interaction region). For CARS-like processes, the intensity of the generated wave is dependent on the square of the integral over the product of the three applied field amplitudes (see Equation (2.28) through (2.30)). Hence, it is advantageous to increase the applied field amplitude density by focusing the beams. Focusing also localizes the interaction region to be a smaller region of the waveguide. This allows the experimenter to avoid imperfections in the waveguide; if an imperfection is found, the applied fields can be crossed in another, more desirable, location.

In the previous calculations, an effective interaction length,  $\ell$ , was defined and, since plane waves were used, the integral over the interaction region is done only over the  $x$ -direction (i.e., the direction transverse to the waveguide plane). Using non-plane waves to deliver the applied fields to the interaction region, the integral over all three dimensions must be taken. Thus, the phase-mismatch parameter,  $\Phi$ , must be redefined as

$$\Phi(\vec{r}) = \Delta\vec{k}(\vec{r}) \cdot \vec{r}, \quad (2.39)$$

where

$$\begin{aligned} \Delta\vec{k}(\vec{r}) \equiv & \bar{k}_1 n(\vec{r}) \sin\theta_{v_1} \cos\alpha_{k_1}(\vec{r}) + \\ & \bar{k}_2 n(\vec{r}) \sin\theta_{v_2} \cos\alpha_{k_2}(\vec{r}) + \\ & \bar{k}_3 n(\vec{r}) \sin\theta_{v_3} \cos\alpha_{k_3}(\vec{r}) + \\ & \bar{k}_4 n(\vec{r}) \sin\theta_{v_4} \cos\alpha_{k_4}(\vec{r}) \quad . \end{aligned} \quad (2.40)$$

Recall the definition for  $n(\vec{r})$  given in (2.10). The radial distance,  $r$ , is best measured from the waveguide coordinate system; it was chosen with some foresight to be the center of the interaction region, that is, where the center of all the beams cross. The azimuthal angles refer to angles of the wavevectors along the profile of each beam. Since the beams are focused, these angles depend on how far from the center of the beam (and the center of the interaction region) the field is being evaluated. Thus, the phase-mismatch parameter is a complicated, nested function of the location from the center of the interaction region. Using non-plane waves also complicates the expression for the amplitude of each applied field. Each beam is dependent not only on the direction transverse to the waveguide plane (i.e., the  $x$ -direction), but also the directions in the plane of the waveguide as well. Hence, the electric field for each wave may be represented as

$$E_{y_i}(\vec{r}) = f_i(x)f_i(y)f_i(z)E_{y_i} = f_i(\vec{r})E_{y_i},$$

where the functional dependence in the  $x$ -direction is determined by the guided mode solution. The dependence of the field amplitude for each wave in the plane of the

waveguide is determined by the focused Gaussian nature of the beam once it has been coupled into the waveguide. With each of these modifications to the simple plane-wave theory given earlier, the *integrated* intensity (i.e., the energy) of the generated wave is

$$U(\omega_4) = \left( \frac{2\pi}{k_4} \right)^2 \left( \frac{\omega_4}{c} \right)^4 I_{\omega_1} I_{\omega_2} I_{\omega_3} \times \left| \int_{-\infty}^{\infty} d\vec{r} \hat{\epsilon}_4 \cdot \chi^{(3)} : \hat{\epsilon}_1 \hat{\epsilon}_2 \hat{\epsilon}_3 f_1(\vec{r}) f_2(\vec{r}) f_3(\vec{r}) e^{i\Delta\vec{k} \cdot \vec{r}} \right|^2, \quad (2.41)$$

where  $I_{\omega_1}$ ,  $I_{\omega_2}$ , and  $I_{\omega_3}$  are the maximum amplitudes of the incident waves at  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ , respectively. If the angular and intensity profiles of each applied beam is known, then this integral can be performed. Next, attention is turned to obtaining expressions for these profiles.

A focused Gaussian beam can be both experimentally and mathematically produced by taking a collimated beam with a Gaussian profile and passing it through a lens. The resulting beam has a wavevector profile and an intensity profile which is given by the Fresnel Diffraction Integral. A Gaussian beam of transverse intensity profile, traveling along a z-axis is specified by

$$\bar{E}_{\text{before}}^{\text{lens}} = \bar{E}_o e^{-(x/x_o)^2} e^{-(y/y_o)^2} e^{-i|k|z}. \quad (2.42)$$

Experimentally, this is the field profile before the collimated laser beam is focused with a

lens. This wave will eventually be coupled into the waveguide. The  $x$ -direction profile of this beam will be projected onto the field amplitude allowed by the mode being coupled into. Hence, the field amplitude profile in the  $x$ -direction is unimportant. As shown by Wlodarczyk and Seshadri [66], a non-plane wave excitation of a grating waveguide coupler does not precisely produce the field profile as expected by the plane wave solutions to the waveguide for an asymmetric waveguide. Moreover, there is a slight focusing effect of a (non-plane wave) beam in the plane of the waveguide when coupled with a grating coupler. It may be noted that no information can be found in the literature concerning the field profile inside a planar waveguide due to a Gaussian beam excitation of a prism coupler. In this analysis, however, slight effects imparted by the coupling of the beam to the waveguide will be ignored. Hence, the Gaussian nature of the beam need only be expressed for the  $y$ -direction:

$$\vec{E}_{\substack{\text{before} \\ \text{lens}}}(y) = \vec{E}_o e^{-(y/y_o)^2}. \quad (2.43)$$

The introduction of a field of a traveling wave through a lens produces a radial (with respect to the beam axis) dependence in the phase of the beam [67]. Hence, the field directly behind the lens appears as

$$\vec{E}_{\substack{\text{after} \\ \text{lens}}}(y) = \vec{E}_{\substack{\text{before} \\ \text{lens}}}(y) e^{-i \frac{|k|}{2f} r}, \quad (2.44)$$

where  $f$  is the focal length of the lens. Propagation of the beam along the  $z$ -axis is described by the Fresnel Diffraction Integral. Assuming the beam is focused by passing

the center of the beam through the center of the lens, the profile of the beam along the  $z$ -axis is described by

$$\bar{E}(y, z, \bar{k}) = \frac{|k| e^{\frac{i|k|y^2}{2z}}}{2\pi iz} \int_{-\infty}^{+\infty} dy_1 \bar{E}_{after}^{lens}(y_1) e^{\frac{i|k|y_1^2}{2z}} e^{-\frac{iy_1 y}{|k|z}}. \quad (2.45)$$

Thus, the functions,  $f_i$ , appearing in (2.41) must be

$$f_i(\bar{r}) = f_i(x) \left| E_i(y, z, \bar{k}) / E_o \right|. \quad (2.46)$$

The absolute value in (2.46) suggests that the phase of the field be included with the function  $f_i$ . With expressions for the profiles of each applied field, (2.46), and a way to calculate the intensity of the generated beam which properly includes the phase-mismatch of crossed, focused Gaussian beams, (2.41); comparisons of large interaction volume experiments versus phase-matched experiments can be made.

## 2.6.2 Results of Phase-Mismatch of WISOM Using Focused Gaussian Beams

Analysis of the generated intensity of collinear beam phase-mismatched conditions where the interaction length is large versus the case where phase-matching is achieved was made. This was accomplished by assuming a 1 mm diameter Gaussian beam (i.e., a 1 mm diameter at  $1/e$  the maximum amplitude), a 1 meter focusing lens, and the results of the previous section. The direction of the wavevector was found for each point in the interaction volume by matrix methods of paraxial optics. A Mathematica notebook, *cars2.ma*, contains the code used to make these calculations. This notebook is located on the enclosed CD-ROM in the *wscars/azimuth1* directory. The nested integrations contained in (2.41) were done numerically in this notebook.



Although numerical calculations were done for many modes, difference frequencies, and waveguides (as enclosed on the CD-ROM), there are two particularly interesting cases. The first case is where the azimuthal angles for zero mismatch are all small, giving a naturally large interaction volume for phase-matched conditions. More commonly, the azimuthal angles are such that the beams are nearly perpendicular to each other making the interaction volume small when the mismatch is zero. This is the second case. To show the variance of intensity versus phase-mismatch, plots were constructed of intensity versus azimuthal angles; the dependence of the azimuthal angles are labeled 'fraction complete'. In the calculation for each plot, the azimuthal angles yielding zero phase-mismatch for the given waveguide, mode combination, and difference frequency were first determined. To generate the plots shown in Figure 2.49 through Figure 2.52, the azimuthal angles are varied from collinear, i.e.,  $[\alpha_1 \equiv 0, \alpha_2 = 0, \alpha_3 = 0, \alpha_4 = 180 \text{ degrees}]$ , to the angles yielding zero phase-mismatch in a smooth way. That is, with  $\alpha_1 = 0$ , angles  $\alpha_2$  and  $\alpha_3$  are gradually changed simultaneously from zero to the phase-matched value by fractions. This way the phase-mismatch also changes gradually. For example, suppose the azimuthal angles that yield zero mismatch are  $[\alpha_1 \equiv 0, \alpha_2 = 20, \alpha_3 = 50, \alpha_4]$ . The point at the far left of one of these intensity versus angle plots is  $[\alpha_1 \equiv 0, \alpha_2 = 0, \alpha_3 = 0, \alpha_4 = 180 \text{ degrees}]$ , that is a 'fraction complete' of zero. This is the collinear-beam geometry. The far right represents 100% complete, that is  $[\alpha_1 \equiv 0, \alpha_2 = 20, \alpha_3 = 50, \alpha_4]$ . Half way into the plot, at a fraction complete of 0.5, the angles are  $[\alpha_1 \equiv 0, \alpha_2 = 10, \alpha_3 = 25, \alpha_4]$ . The output angle,  $\alpha_4$  is found by minimizing the phase-mismatch once each of the input angles have been chosen.

An example of the case is where the azimuthal angles for zero mismatch are all small is shown in Figure 2.49 and Figure 2.50. Here there is a large interaction volume for phase-matched conditions. Figure 2.49 shows the variations in intensity due to phase-mismatch of a waveguide with a 1.85 index and 800 nm thickness using mode combination  $[2,2,0,1]$

at a difference frequency of 1200 wavenumbers. With this difference frequency and the choice of  $\omega_1 = \omega_3 = 560$  nm, the wavelength of the beam at  $\omega_2$  is 639 nm, making the output beam of wavelength 409 nm. This choice of WISOM parameters gives the phase-matching azimuthal angles of  $[\alpha_1 \equiv 0, \alpha_2 = -1.44, \alpha_3 = 44.9, \alpha_4 = 219.1]$  degrees. This calculation is reflected in the text file *cars6m80.xl* on the enclosed CD-ROM. In this plot, it is clear that when the applied beams are collinear, the phase-mismatch is large and the generated CARS-like signal is low in intensity (as shown on the left-hand side of Figure 2.49). This is relative to the large intensity obtained when the process is phase-matched as shown on the right-hand side of the plot. For this example, the ratio of intensities between using the phase-matched angles versus collinear beams is 2.97. Clearly, when the phase-matching azimuthal angles are small, it is advantageous to phase-match the OM process rather than to maximize the interaction volume by using collinear beams. It is noteworthy to point out the *sinc*-like behaviour of the intensity on phase-mismatch (i.e., azimuthal angle 'fraction complete'). To better show this dependence, the plot shown in Figure 2.49 has been amplified in Figure 2.50. Figure 2.50 clearly shows a *sinc*-like behaviour; this is modulated by a high-frequency  $\cos^2$ -like dependence on the phase-mismatch. Due to aliasing in the calculations, this  $\cos^2$ -like behaviour is not well reproduced. This is despite the step size of 1/500; a step size which gives a calculational time of less than 18 hours. When the WISOM process is phase-matched with a nearly collinear azimuthal angles, the large interaction volume is appreciable. It is not surprising that, intensity-wise, in this situation it is most effective to phase-match the WISOM experiment rather than try of maximize the interaction volume by making the beams collinear. Each calculation is normalized so that all calculations can be directly compared. The normalization is done by assuming the phase-matched situation is that in which all beams are collinear, the intensity of this case is found, and used as the normalization factor.

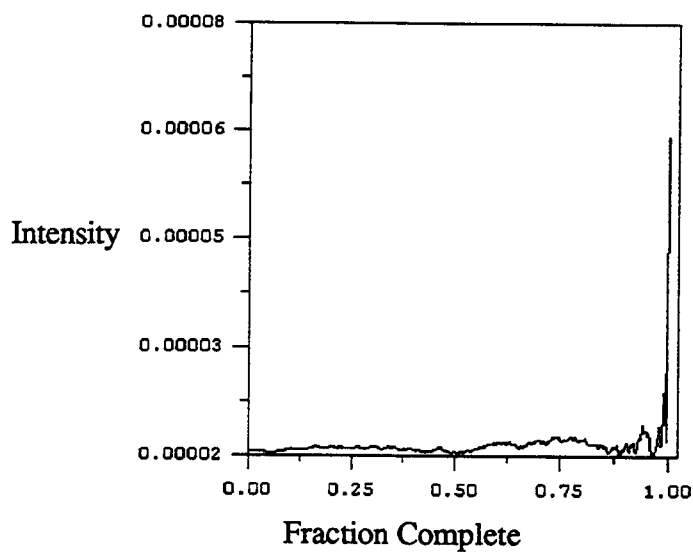


Figure 2.49: A plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at  $\Delta = 1200 \text{ cm}^{-1}$ .

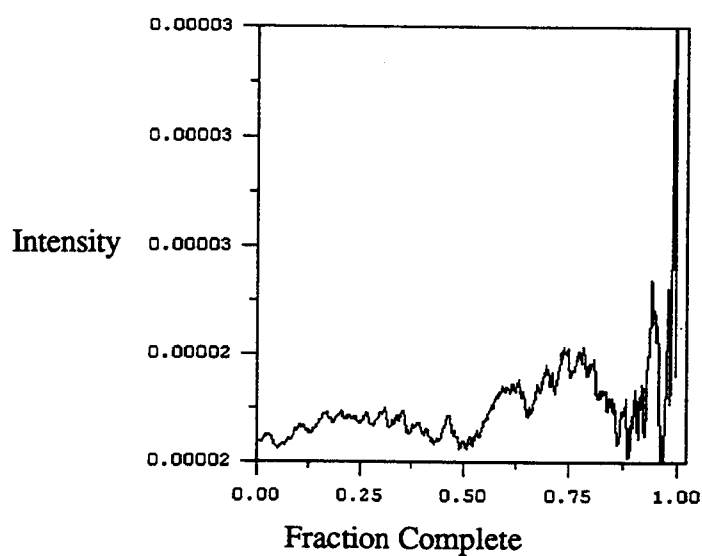


Figure 2.50: An expanded plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at  $\Delta = 1200 \text{ cm}^{-1}$ .

It is more common that the azimuthal angles giving a mismatch of zero are nearly right angles; this makes the interaction volume small compared to the collinear application of the input beams. Despite the shorter interaction length, the phase-matched case generally provides signals two to three times larger. A typical example is shown in Figure 2.51. Here a plot of intensity versus phase-mismatch (or ‘fraction complete’) for a 1.85 index waveguide with an 800 nm thickness using mode combination [2,2,0,1] with a difference frequency of 2000 wavenumbers is shown. This makes the wavelengths of the beams 560, 631, 560, and 504 nm for  $\omega_1$  through  $\omega_4$ , respectively. The phase-matching azimuthal angle for this set of WISOM parameters is [ $\alpha_1 \equiv 0$ ,  $\alpha_2 = -90.0$ ,  $\alpha_3 = 80.6$ ,  $\alpha_4 = 191.8$ ] degrees. Text file *cars6k80.xl* contains the results of this calculation. The  $\cos^2$ -like behaviour of the intensity on phase-mismatch is very apparent; it makes the interpretation of the overall phase-match dependence difficult. Hence, a windowed average of the plot in Figure 2.51 was taken in order to de-emphasize this behaviour. The averaging window, which is moved across the plot, was taken to have a width ten times the highest frequency in Figure 2.51; that is, 0.15 ‘fraction complete’. Figure 2.52 shows this averaged plot. The effects of a large interaction volume on the intensity are seen on the left-hand side of the plot. When the beams are collinear, the mixed signal intensity is large. As the angles are increased away from collinearity, the signal drops. On the far right-hand side, the intensity is greater when the process is phase-matched. This is despite the decreased interaction volume. When only plane waves are considered, a *sinc* dependence on the phase-mismatch is predicted. A *sinc*-like dependence is vaguely apparent near the phase-matched side of the plot. The intensity ratio of the phase-matched, small interaction region condition (at 1.00 ‘fraction complete’), versus the collinear case (at 0.00 ‘fraction complete’) is 2.8. Apparently, phase-matching gives a larger overall intensity, smaller interaction volume notwithstanding.

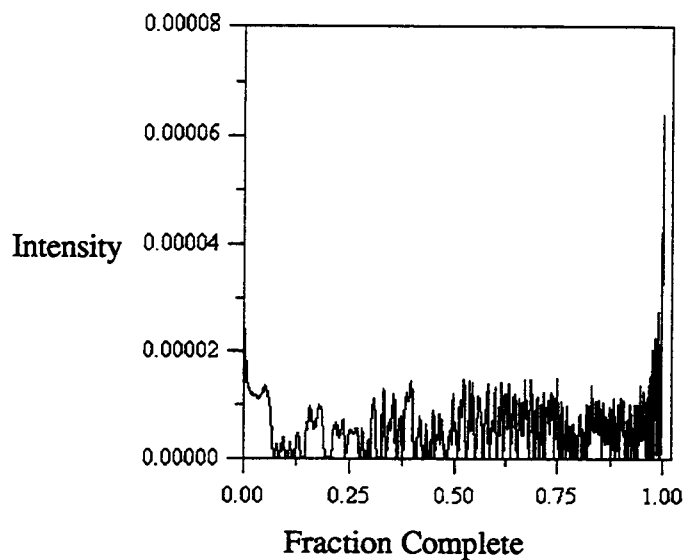


Figure 2.51: A plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at  $\Delta = 2000 \text{ cm}^{-1}$ .

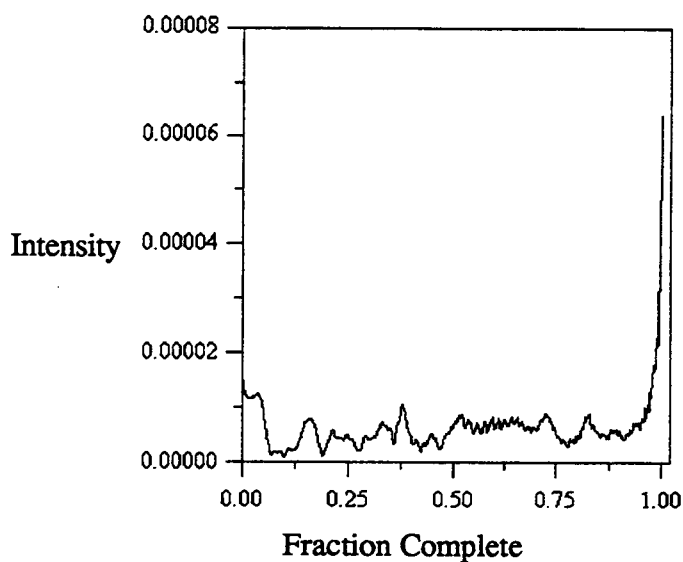


Figure 2.52: An averaged plot of the normalized intensity versus azimuthal angles for a 1.85 index, 800 nm thick waveguide with mode combination [2,2,0,1] at  $\Delta = 2000 \text{ cm}^{-1}$ .

Despite the shorter interaction length, the phase-matched case generally provides signals one to three times larger than the collinear beam case. It was assumed that any dispersion in the cover and in the film layers were identical. As argued in Chapter 1, phase-matching may be possible in the surface region while the bulk (i.e., film) region is left with a significant phase-mismatch. This occurs when the dispersions of the two layers are different. Calculations (see *cars2a.ma*) show that for a dispersion of one part in  $10^{-3}$  for the surface and twice that for the film, typically, there is an additional enhancement in the surface-to-film ratio of  $10^3$  to  $10^4$ . This is a stunning and potentially useful result.

## 2.7 Couplers for Four-Beam Mode Combinations

The theoretical analysis shows that four-beam mode combinations allow for phase-matching, and at the same time, provide broader tuning ranges. These analyses show that there several advantages in using these four-beam mode combinations. Unfortunately, using four-beam mode combinations increases the experimental difficulty. Since the waves are all non-degenerate in wavevector, the three beams must be coupled into the waveguide. The generated waves, those comprising the fourth beam, must also be coupled out of the waveguide and prepared for detection. Calculations show that to achieve phase-matching, the azimuthal angle between each beam typically needs to be approximately 90 degrees. To take advantage of phase-matching in four-beam mode combinations, four couplers are needed to couple four beams to the waveguide. It should be noted that it is very difficult to construct four different couplers on a single waveguide. Hence, four separate couplers are required as shown in Figure 2.53.

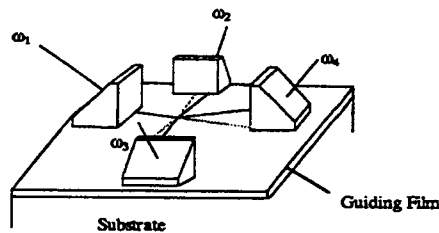


Figure 2.53: A WSCARS coupling scheme for four-beam mode combinations.

Using prism couplers, it is hard enough to get two functional couplers on a waveguide, let alone four. Thus, some efforts were made to improve couplers and the construction of couplers on planar waveguides for the purpose of WISOM experiments utilizing the four-beam mode combinations. This section discusses several attempts to improve the ability to couple separate beams into the asymmetric waveguides. First, achromatic waveguide couplers are discussed. This modified prism coupler simplifies the use of prism couplers by eliminating the need to re-adjust the (polar) angle as the frequency of the beam is changed. This approach, however, cannot solve the problem of the variability of the coupling efficiency inherent in prism couplers. The other approach, considered at some length, is the use of integrated grating couplers as shown in Figure 2.57. Attempts were made to integrate grating couplers into waveguides with limited success. Experiments with constructing, reproducibly, these integrated couplers are discussed. Fabrication techniques for line gratings are covered. Again, the problem of constructing four such integrated couplers proved challenging. One proposed solution is to integrate a nonlinear medium into the substrate material and induce a phase grating which could be used to couple the WISOM beams into the waveguide. This solution is briefly discussed. Suggestions on the use of diffractive optics to construct ring grating couplers were explicitly investigated. Analytical and numerical predictions concerning the use of diffractive optics completes this section.

### 2.7.1 Achromatic Prism Couplers

To improve the usefulness of prism couplers, achromatic couplers can be considered. One problem of any coupler is the constant readjustments that must be made in the coupling (i.e., polar) angle as the frequency of the waves are changed. An achromatic coupler minimizes these adjustments, thereby simplifying the WISOM experiment. Such couplers were proposed by Spaulding and Morris [68]. One such coupler, a hybrid prism-grating coupler is shown in Figure 2.54. The prism-grating coupler, made from Schott

SF1 glass, when placed on a waveguide of a Corning 7059 film on a Pyrex substrate was shown to provide achromatic operation over a 33 nm tuning range centered around 605 nm. Over this range the coupling angle tolerance was  $\pm 0.005$  degrees. The compensating grating spacing was 352 nm when the prism angle was 62 degrees. Over a broad tuning range, the sensitivity to coupling angle is reduced. This is shown in Figure 2.55 where the error in (polar) coupling angle is plotted over a range of 570 to 640 nm; the hybrid coupler can be compared with the traditional prism coupler. Figure 2.56 compares the coupling efficiency of a prism coupler with this hybrid prism-grating structure. Using their theoretical expressions, such a coupler build from  $\text{SrTiO}_3$  prism of angle 30 degrees would give a tuning range of 314 nm for a waveguide as has been considered in the previous sections. Use of achromatic couplers does not, however, solve the problems of the variability of coupling efficiency. This the prism coupler's downfall; prism couplers must be pressed onto the film of the waveguide such that the bottom face of the prism is within a wavelength of the film. Irregularities in the flatness or index of either the waveguide or prism cause the coupling efficiency to vary over the bottom face of the prism. In practice, the experimenter clamps the prism to the waveguide and searches for 'magic locations' along the prism face where the coupling efficiency is favourable. Frequently, no 'magic locations' are found and prism must be removed and re-applied to the waveguide. This may have the undesirable effect of damaging the waveguide.

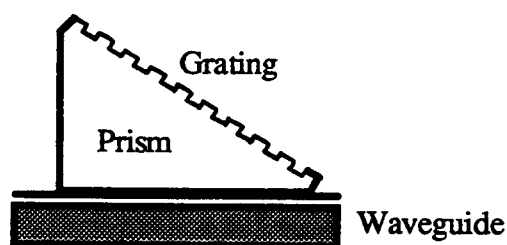


Figure 2.54: Spaulding and Morris's prism-grating coupler.



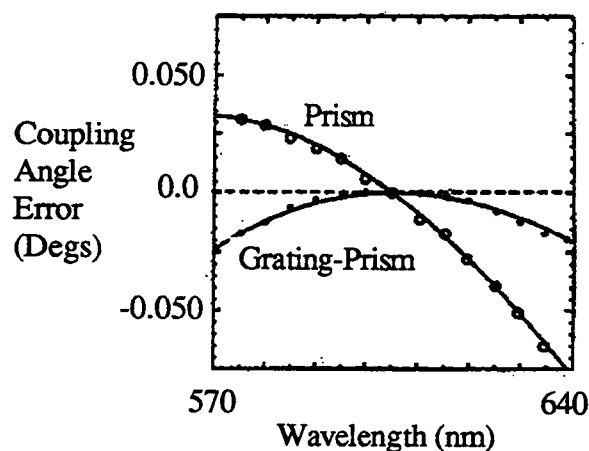


Figure 2.55: Angular sensitivity to tuning for Spaulding and Morris's prism-grating coupler.

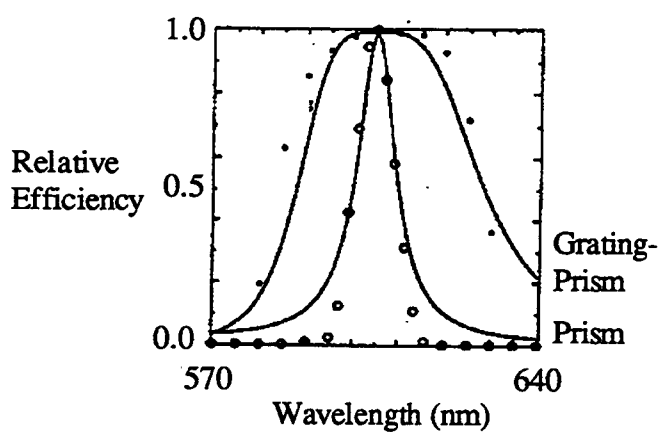


Figure 2.56: Coupling efficiency of Spaulding and Morris's prism-grating coupler.

### 2.7.2 Integrated Grating Couplers

Another approach to coupling and decoupling four separate beams into a waveguide is the use of integrated grating couplers as shown in Figure 2.57. Such grating couplers

have been extensively investigated and are well characterized. (See, for example, [66] and [70].) For purposes of CARS-WISOM work, four such grating couplers per waveguide must be fabricated, and ideally, have spacings of 450 to 500 nm. Clearly, fabrication techniques must provide good reproducibility in order to construct this many gratings per waveguide. Typically, gratings they are constructed on the film material. However, there are the advantages of resilience to damage and dirt when constructing the gratings on the substrate before the film is applied. This was the direction taken in this research.

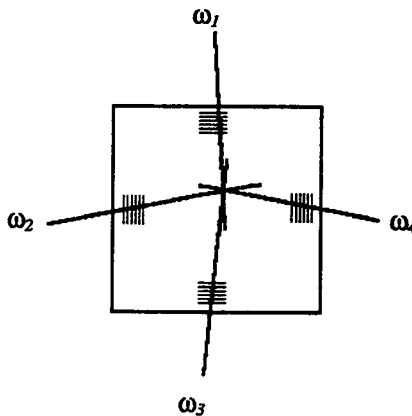


Figure 2.57: A four coupler scheme for planar waveguides using integrated grating couplers.

Experimental attempts were made to integrate line grating couplers into waveguides with limited success. Experiments with constructing reproducible gratings on substrates were performed. Fabrication techniques for line gratings were followed according to previously outlined methods [69], [70], [71]. The process involves spin-coating a UV-sensitive photoresist (Shipley Microposit S1400-31) to a thickness of 800 to 1200 nm on a microscope slide substrate. Preparations to spinning included a complete cleaning of the substrate in acetone, methanol, followed by de-ionized water. Samples were heated to 200 Celsius for 1 hour. The substrate was then transferred to a small clean environment equipped with a sub-micron air filter. This 'clean box' was constructed to house the

spinner and to provide dust-free work surfaces and sample storage. Several drops of hexamethyldisilazane (HMDS) were applied to the substrate after it was affixed to the spinner and was spun dry. This prepares the surface of the substrate allowing the photoresist to better adhere. The HMDS sample was then baked at 200 Celsius for 1 hour. Photoresist, diluted 2:1 with filtered xylene was spun onto the substrate. The diluted photoresist was filtered through a 0.5 micron filter prior to spinning. Spinning parameters were adjusted to give photoresist thicknesses of 800 to 1200 nm. The resist was allowed to dry overnight in the light-tight clean box. Exposure of a cosine-squared intensity pattern was accomplished using a Lloyd's Mirror Interferometer (LMI) as described in [69]. A UV argon-ion laser (356 nm) was spatially filtered using a 40X microscope objective and a 15 micron pinhole. The beam was collimated to a 7 mm diameter using a 25 mm fused silica lens. This beam was applied to the LMI. Exposures of 10 to 15 seconds with an integrated beam power of 45 mW were applied to the photoresist. The exposed photoresist was placed in a holder-dish and surrounded by developer. Developing was done using Shipley Microposit Developer diluted 3:1 with DI water. A monitoring system [72] was used to guide the development process. This system consisted of a He-Ne laser with an expanded diameter beam of roughly 10 mm. The intensity of the diffracted beam is monitored and development is stopped by flushing with DI water when developing is optimal. Gratings spacings of various sizes were produced. Developed grating spacings (in the photoresist) were measured to be 477 to 2785 nm using LMI angles (as measured from the normal of the substrate) of 3 to 15 degrees. Once the procedure had been established, gratings in the photoresist were found to be very reproducible in both spacing and diffraction efficiency. Four substrates containing 24 gratings each were produced. All gratings were shown to have identical spacings and efficiency. Out of the 96 total gratings developed on 6 substrates, only two were considered unusable. The next step was to etch the gratings into the substrate using the photoresist as a mask. This was done using 10% HF. The same holder-dish and monitoring system was used during the etching process. Commonly, the photoresist was found to peel away from the substrate before the etching was complete. The success rate

for producing gratings in this manner was very poor; only one in twenty gratings were etched into the substrate, and those gratings have less than a 1% diffraction efficiency into the first order. Later, it was suggested that the sample be periodically removed from the dish, rinsed with DI water, and reinserted into the etching acid. During the HF etching process, a chemical barrier forms on the exposed glass surface preventing effective etching from occurring. Periodic washing removes this barrier. Without removing this barrier, the HF eventually dissolves the photoresist and destroys the grating. Other methods of etching could also be employed including ion milling.

The problem of constructing four such integrated couplers per waveguide proved to be too challenging. One disadvantage to grating couplers is that, once constructed, they cannot be moved. The placement of the gratings on the waveguides must be determined in advance. As seen in Section 2.6.2, the ‘crossing angle’ for minimum phase-mismatch changes for a given mode combination as the difference frequency is tuned. Either the experimenter must take into account the effects of intensity due to mismatches or adjust the azimuthal angles such that zero mismatch is achieved. In the latter case, the beams may not be incident on the grating perpendicular with respect to the lines in the grating coupler. (In the case of a prism coupler, the beams may not be perpendicular to the angle in the plane of the waveguide). Not diffracting off the grating at a 90 degree angle affects the coupling efficiency and potentially misshapes the beam profile [66]. Moreover, as the crossing angle is changed, the interaction region is moved along the waveguide; this may be undesirable. Hence, in a following section, ring grating couplers are considered.

### 2.7.3 Integrated Nonlinear Grating Couplers

Before considering ring gratings, the possibility of inducing gratings in a photorefractive medium embedded in the waveguide is considered. Instead of constructing a fixed grating in the substrate, a photorefractive material such as lithium niobate or barium titanate would be grown onto the film of the waveguide. A sinusoidal

grating would be induced by applying two interfering beams as shown in Figure 2.58. The interfering ‘pump’ beams induce a spatial dependence in the index of refraction of the photorefractive coupler. This grating formation is described in [73]. The index modulation is then used as an adjustable, active diffraction grating. The WISOM beams would then be coupled into the waveguide using these active gratings. Although the efficiency of these gratings is typically not expected to be exceptional, the ability to alter the orientation and spacing of the gratings would be an attractive advantage.

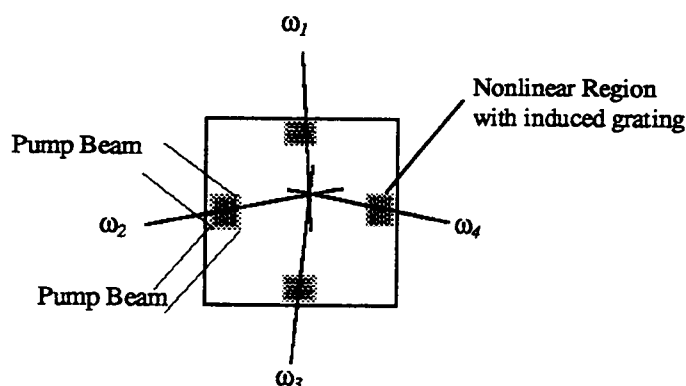


Figure 2.58: A four coupler scheme using induced grating couplers in embedded nonlinear regions.

#### 2.7.4 Ring Gratings

Since the azimuthal angles for minimum phase-matching are very different for different mode combinations and difference frequencies, another alternative was sought to couple the applied beams to the waveguides. It is suggested that ring gratings be constructed in place of linear gratings for coupling into the waveguides. This section outlines this suggestion. Although no gratings of this type have been demonstrated, this section details their proposed construction methods.

Ring gratings could be constructed on the substrate of a waveguide prior to applying

the film or fabricated into the film after application. The ring grating would surround the interaction region with an clear diameter of 3 mm as shown in Figure 2.59. The grating spacing, that is the spacing between rings, could be constant or could be slowly chirped to increase the tuning range. In either case, the grating spacing is optimal for CARS-WISOM when the spacing is in the range 450 to 500 nm. Fabricating these structures could be done using the same methods for making linear gratings as described in Section 2.7.2. The interference pattern used to expose the ring grating structure (in the photoresist) may be created in two different ways. First, is the use of the diffraction of a circular aperture. The diffraction pattern of a circular aperture is a  $J_0$  function; far from the center of the pattern, the spacing of the fringes is constant. Note that if a spacing chirp is desired, the center of the pattern may be used since the fringe spacing is chirped; the  $J_0$  maxima are further apart (in radius) near the center of the pattern. The second method involves the use of diffractive optics to generate the ring structure directly.

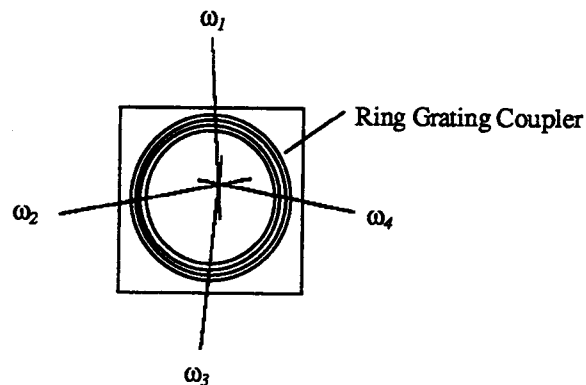


Figure 2.59: A ring grating waveguide coupler.

Creating the ring grating exposure pattern via a circular aperture would involve masking a  $J_0$  diffraction pattern. The apparatus for generating the pattern is shown in Figure 2.60. As with exposing a linear grating, a collimated spatially filtered UV laser is

used a source of plane waves. A circular aperture forms the  $J_0$  diffraction pattern. An auxiliary expanding lens may be used to expand the pattern and an opaque disk is used to mask off the center portion of the  $J_0$  diffraction pattern leaving only the ring grating intensity pattern. This pattern is represented in intensity cross-section in Figure 2.60. This pattern is used to expose the photoresist spun onto the substrate.

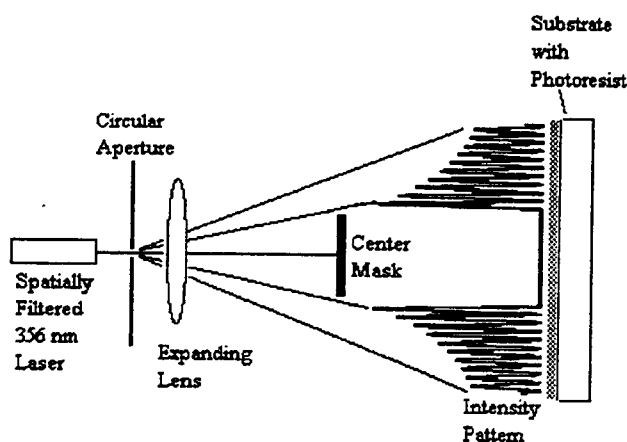


Figure 2.60: Creating a ring grating exposure pattern using a circular aperture.

Using a 356 nm source, the circular aperture should be 1.9 mm which would be placed 5 mm from the photoresist. Masking the center 3 mm of the resulting  $J_0$  diffraction pattern at the photoresist will produce the clear aperture, while leaving a 468 nm ring grating. This grating spacing varies only 2% over 5000 rulings. The exposure time, based on the results and exposure intensities in Section 2.7.2, is many hours. A longer exposure is required since most of the light is in the first few orders of the diffraction pattern which are blocked to create the clear aperture. The energy per fringe in the fringe pattern used in the fabrication of linear gratings (using the Lloyd's mirror interferometer) can be calculated. Likewise, the energy per fringe using the blocked  $J_0$  diffraction pattern method for creating ring gratings can be calculated. Hence,

a direct comparison can be made between the exposure times of the linear gratings and ring gratings using the blocked  $J_0$  fabrication method. The exposure per fringe for the linear gratings using the 350 mW UV source over a 10 second exposure is 74 mJ. Using the same source, the power illuminating one fringe of the ring grating, on average, is 580  $\mu$ W. Thus, the same 74 mJ exposure per fringe of the ring grating would require over 13 hours using this fabrication method. These calculations are outlined in the file */wscars/grating/ring4.ma*.

Such a long exposure time is a great disadvantage since vibrations of a few nanometers over that time scale are likely; these vibrations would destroy the exposed image, and hence, the grating pattern. Another alternative is the construction of a diffractive optic to create the ring pattern directly. The advantage to a diffractive optic is that nearly all the intensity from the source is directed into a pre-determined pattern; exposure times comparable with linear gratings would be possible. Essentially, a diffractive optic is an intensity mask that is placed in front of a (Gaussian profile) laser beam. It is carefully designed as to produce the requested pattern. The intensity profile of the diffractive optic used to create the intended pattern can be obtained using the results of Fourier optics. Recall the Fresnel Integral,

$$E(x_o, y_o) = \frac{ke^{ikz}}{2\pi iz} \int_{-\infty}^{\infty} dx_1 \, dy_1 \, E(x_1, y_1) e^{i\frac{k}{2z}[(x_o - x_1)^2 + (y_o - y_1)^2]}. \quad (2.47)$$

Knowing the field everywhere in the plane  $x_1 - y_1$  perpendicular to the optical axis, the Fresnel Integral calculates the field at a point  $(x_o, y_o)$  in a plane down the optical axis.



Coherent, monochromatic light is assumed. In the limit where the two planes are far apart compared with the wavelength of light, that is the Fraunhofer limit, the Fresnel Integral becomes a Fourier transform:

$$E(x_o, y_o) = \frac{ke^{ikz}}{2\pi iz} e^{i\frac{k}{2z}(x_o^2 + y_o^2)} \int_{-\infty}^{\infty} dx_1 dy_1 E(x_1, y_1) e^{-i\frac{k}{2z}[x_o x_1 + y_o y_1]}. \quad (2.48)$$

When the fields are cylindrically symmetric, the Fresnel Integral in the Fraunhofer limit becomes a Hankel transform of zero order. It takes the form of

$$E(r_o) = \frac{ke^{ikz}}{iz} e^{i\frac{k}{2z}r_o^2} \int_{-\infty}^{\infty} dr_1 E(r_1) r_1 J_0\left(\frac{k}{z} r_1 r_o\right). \quad (2.49)$$

Apparently, a diffractive optic can be designed by obtained by taking the inverse Hankel transform of the desired pattern. In this case, the desired concentric circular pattern was chosen with no intensity inside a radius of 1.5 mm and a square-wave function for radii exceeding 1.5 mm. The equivalent periodicity for the square wave function was chosen as 450 nm. After 5000 fringes (or rings) the square wave function was terminated. Thus, the pattern would have 5000 concentric ring fringes with the first fringe having a radius of 1.5 mm; all the light would exist in this pattern and no where else. The Mathematica notebook */wscars/grating/ring4b.ma* was written to find the diffractive optic that would create this pattern. The intensity profile for the diffractive optic that creates this ring pattern is shown in Figure 2.61. This figure shows the relative transparency of a transmission mask across its center; the pattern is circularly symmetric about the vertical axis shown in the plot. Construction of this profile into a usable transmission mask could be done using shaded computer images transferred to a film recorder and subsequently

reduced photographically. This method is used in creating diffractive optics, however, optics of this scale and accuracy are not often fabricated.

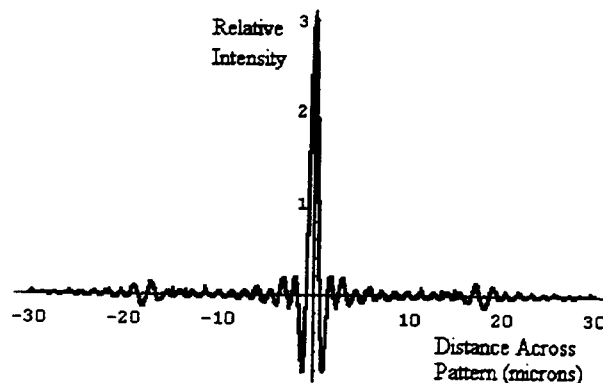


Figure 2.61: Intensity profile of a diffractive optic for creating a ring grating exposure pattern.

These suggestions for coupling into planar waveguides are very much academic. Further research and experimentation into the fabrication of alternative grating couplers such as ring gratings and integrated nonlinear grating couplers is needed. Unanswered questions include: what are the required exposures, what are the efficiencies, how robust, and with what precision can these grating patterns be constructed.

## 2.8 Conclusions

Surface-specific experiments can be done in a planar waveguide utilizing an engineered interference condition. This technique is referred to in this research as WISOM. To gauge the surface-specificity of a WISOM process, a theoretical development for planar asymmetric step index waveguides using plane waves was introduced. This theoretical treatment yielded expressions for the field amplitude and phase of an induced guided wave in terms of the incident plane waves parameters, the linear indices, and nonlinear susceptibilities of the cover, film, and substrate. Interpretative tools were introduced to quantify the surface-specificity. The theoretical treatment was then used to perform

numerical calculations which were used to predict favourable WISOM conditions using experimentally realistic parameters. The numerical analysis comprises a data base which may be used by future researchers. Six mode combinations were found to be noteworthy, two of which have not been mentioned in the literature. Some had exceptional surface-specificity, some were found to have large tuning ranges, and one had small phase-matching angles. Both three-beam and four-beam mode combinations were analyzed.

Phase-matching inside a waveguide had never been investigated before. It was found that the ability to phase-match in the waveguide is not uncommon. However, phase-matching often comes at the expense of a large interaction volume. An experimentally appropriate theoretical development was constructed to evaluate whether a large interaction volume or a phase-matched process is most desirable. The development introduced an expression for the integrated intensity of an OM process in a waveguide. A numerical analysis ensued. It provided the conclusion that, in defiance of the shorter interaction length, phase-matched geometries generally provide signals two to three times larger than collinear beam geometries. It was found that there are benefits to employing phase-matching and utilizing the four-beam mode combinations; however, coupling four beams to a waveguide is fraught with problems. Efforts were made to improve coupler design and construction for asymmetric planar waveguides. Coupling schemes considered include: achromatic prism couplers, integrated line grating couplers, coupling gratings induced in an integrated photorefractive material, and ring gratings couplers produced by masked circular diffraction patterns or diffractive optics.

The final conclusions are that WISOM is fraught with experimental difficulties and provides limited surface-specificity. Clearly, WISOM can only be achieved on the surfaces of thin dielectric films that can be reproducibly grown on dielectric substrates. These thin films must have relatively high indices of refraction capable of supporting guided modes. If prism couplers are to be employed, these films must be quite robust. Scanning through the resonances of the surface using the WISOM technique may require a

series of waveguides due to the narrow tuning ranges of some waveguides and mode combinations. Breaking the beam degeneracy of traditional experiments and using four-beam mode combinations or phase-matching is a way of increasing signal levels. When there is a disparity in dispersion between the film and the surface, non-degenerate experiments can be very effective in providing surface-specificity. Unfortunately, these non-degenerate experiments produce the added aggravation of having to couple many beams to the waveguide. Despite the disadvantages and complexities of the WISOM technique, it can be a valuable surface-specific characterization technique. However, as Chapter 1 hints, there may be more straightforward techniques that provide surface-specificity.

### 3. REFLECTION GEOMETRY INTERFACE-SPECIFIC OPTICAL MIXING (RISOM)

#### 3.1 Introduction

Can vibrational and electronic surface resonances be detected on any flat surface? This has been the central question of this research. Perhaps a more specific statement of this question is: How can the surface-to-bulk signal ratio be maximized in ISOM experiments? Since there is a small surface signal riding atop a large bulk signal, how can the typically poor signal-to-noise ratio be maximized? Chapter 2 took the approach of previous researchers: using FWM in waveguide geometry to maximize the surface-to-bulk signal ratio. One conclusion was that to experimentally scan through surface resonances, several waveguides would be required since a single waveguide can provide favourable conditions only over a limited range of frequencies. A second conclusion from Chapter 2 is that use of phase-matching in a waveguide can be advantageous, but the configuration of the incident fields is far more complex than the traditional collinear beams experiment. Finally, Chapter 2 concludes that WISOM limits the study of surfaces of dielectric materials to those for which favourable waveguides can be constructed. This is a serious restriction considering the number of dielectric materials for which good films can be grown of the right thickness and indices of refraction. There are many more surfaces of interest than this limited set.

Chapter 2 answered the original question for a very limited case. Borrowing from the wisdom of Talmudic philosophy, it must be suggested that the right question is not being asked. Attempting to ask a more specific question, a new query was posed: Can simple reflection geometry be employed in an optical mixing process and still be used to detect a surface component to a signal that is composed mostly of bulk contributions? A reflective geometry experiment is straightforward and phase-matching is easy to achieve; signals are

typically large and there are no apparent limitations on the type of surface that can be studied. A typical three-wave mixing process in reflection geometry experiment is shown in Figure 3.1. The inset provides a microscopic perspective. There, a surface region (the lightly shaded region) and a semi-infinite bulk (the unshaded region) are shown. The two applied fields with wavevectors  $k_1$  and  $k_2$  mix together in the ‘interaction region’ (the heavily shaded region) and generate the field at  $k_3$ . Note that the interaction region includes volumes in both the surface and bulk regions. Thus, in general, light is generated from both regions.

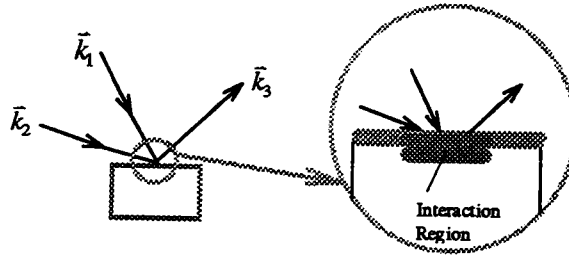


Figure 3.1: Diagram of a typical TWM experiment in reflection geometry.

Contrasted with the waveguide geometry, there is no built-in interference condition in RISOM. Another difference is that there are no restrictions, other than the surface being flat on the scale of one wavelength, placed on RISOM geometry. Since RISOM does not require a particular structure, there is no limitation on the types of surfaces that can be studied. In reflective geometry, where experiments are very tractable, there is complete freedom in choosing the applied field wavevectors and polarizations. Not unlike reflections in the linear case, nonlinear optical processes in reflection geometry have angular and polarization dependencies. Based on these dependencies, engineering of the applied fields is the path to favourable surface-to-bulk signal ratios in RISOM.

Exactly why should there be friendly signal-to-noise ratios in RISOM experiments? To elucidate, consider a semi-infinite medium comprised of layers of dipoles free to

oscillate. This is pictured in Figure 3.2. In this framework, the top layer -- the surface layer (the gray layer in Figure 3.2) -- might have different dipoles (in strength and orientation) than the rest of the 'bulk'. Experiments indicate this is so, as excellently reviewed by Corn and Higgins [9]. Now suppose the dipoles are forced to oscillate by applying incident harmonic fields. The process is linear when the dipoles oscillate with the same frequency as the applied fields, that is, a *linear* reflection occurs. The angular and polarization dependencies are described by the familiar (linear) Fresnel equations. By analogy, when the process is nonlinear, the dipoles oscillate with (in general) a different frequency than the applied fields. Since the strength and orientation of the dipoles with relation to the applied and generated (reflected) fields is important, there are *nonlinear* Fresnel equations that describe the angular and polarization dependencies; these will be different from the linear Fresnel equations. Also, since the dipoles are different at the surface, there will be different nonlinear Fresnel equations for the surface and bulk. This means the angular and, perhaps polarization, behaviour is different for the two regions. Based on this simple argument, it seems that there might be a way to separate the surface and bulk signals in a nonlinear optical process that employs reflection geometry.

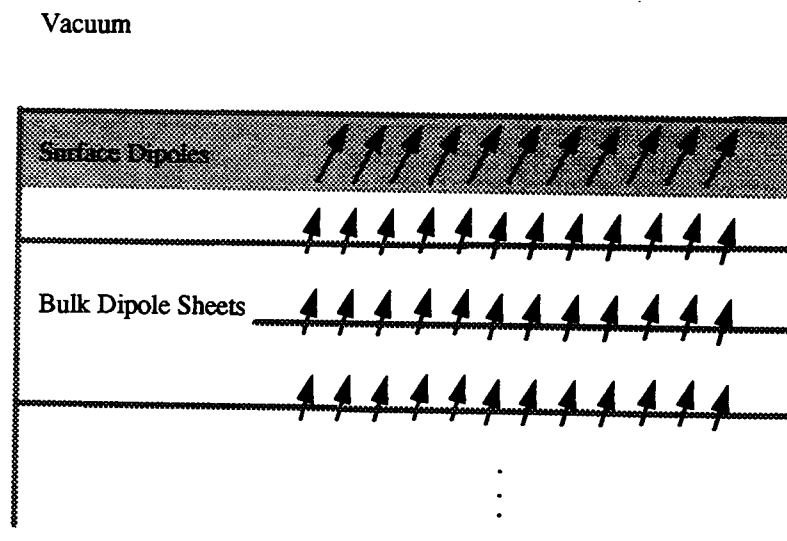


Figure 3.2: A semi-infinite medium comprised of layers of dipoles.

The simple dipole sheet model used above is undoubtedly the most appropriate approach to developing RISOM theory that predicts the angular and polarization dependencies of bulk and surface signals. Such a model has been used to calculate the microscopic local fields *inside* a semi-infinite dipole structure where all the dipoles were identical [22]. It proved very successful, but involved. Rather than do this time-consuming and arduous calculation, a less precise but much simpler macroscopic model was adopted. In a series of papers, Bloembergen and co-workers determined the theoretical equations that govern optical mixing (particularly SHG) in reflection geometry [23], [24], [25]. This *macroscopic* approach assumed optical mixing occurred within a thin slab of nonlinearly active material. An interface between this layer and a semi-infinite bulk having only linear properties (but different than that of the surface slab) provides a mechanism for the generated light to reflect and return to the vacuum. This structure is pictured in Figure 3.3. Note that without this internal boundary, the generated light could never return to the vacuum. This is a consequence of assuming that the new fields are generated by a *macroscopic* nonlinear polarization induced in the surface layer. What does it mean that the linear indices of bulk and surface are different? From the (linear) Fresnel coefficients, the reflection from this interface is stronger as the difference gets larger. Thus, the surface-to-bulk index ratio is a scaling factor in the amplitude of the generated field as measured in the vacuum.



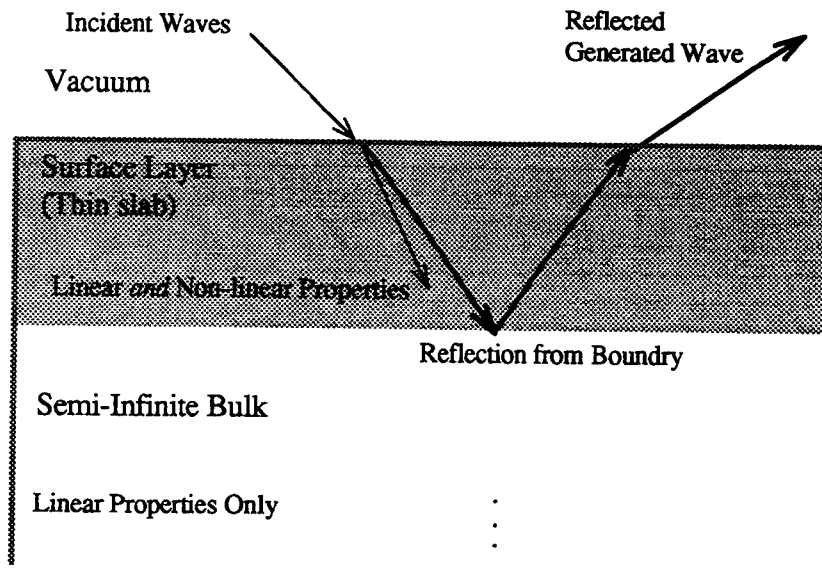


Figure 3.3: Optical mixing in a thin slab of nonlinearly active material.

What justification is there for using a macroscopic model for a process that has microscopic origins? It is indisputable that the motive of nonlinear processes (optical mixing) is microscopic. Analysis by Ju [22] of sheet dipole model (with identical dipoles) shows the classic linear Fresnel coefficients and Bloembergen's macroscopic nonlinear thin slab model (both derived from a macroscopic approach) to yield results that are not exact, but exhibit the main attributes of a true microscopic model. His work suggests that use of a macroscopic model can furnish, to good approximation, linear and nonlinear reflection coefficients. Use of a macroscopic model in this work will be used to guide future work; at some juncture, the microscopic calculations must be done. Embarking on the microscopic calculation without prior direction would be an inefficient use of resources.

Since the goal of RISOM is to achieve favourable surface-to-bulk signal ratios, any successful model must be able to describe signals generated from a surface region *and* the bulk. With such a model, the surface-to-bulk signal ratio can be constructed and the parameter space explored for RISOM conditions. Macroscopic models investigated to

date, including the Bloembergen model, describe signals that originate from a thin surface slab; no contributions from the bulk are included. Recalling the Bloembergen model structure, there were no nonlinear properties included for the semi-infinite bulk. It is insufficient to simply include nonlinear properties for the bulk. Just as the light generated from the thin surface slab needs an interface to reflect off in order to return to the vacuum (recall Figure 3.3), so does the light generated from the bulk. Hence, the following structure, as shown in Figure 3.4, was adopted.

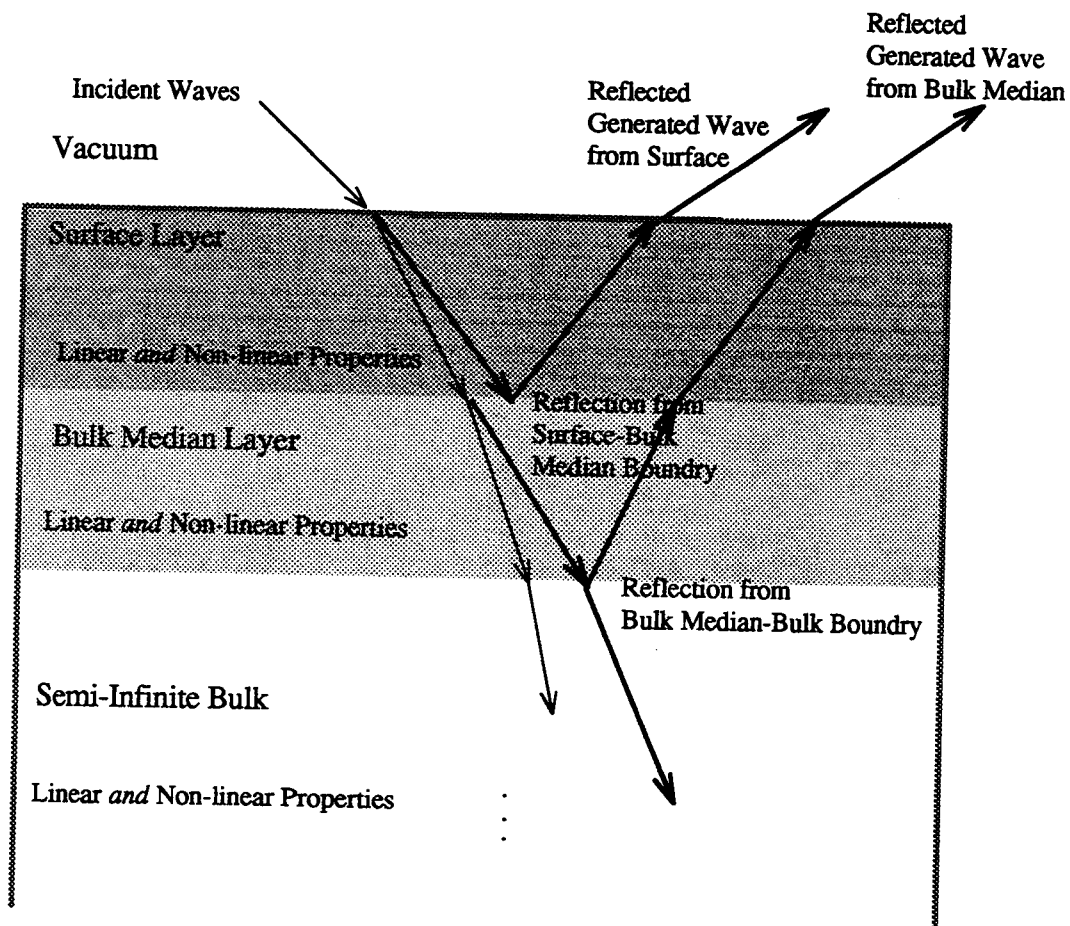


Figure 3.4: A structure that gives reflected optical mixing signals for surface and bulk regions.

Extending Bloembergen's model, an additional layer is added between the surface slab and the bulk. This is referred to as the *bulk median*. This layer has the nonlinear properties of the bulk which are, in general, different from the surface. The linear index of refraction of the bulk median, however, is made different than the semi-infinite bulk. This new interface establishes a mechanism for the light generated in the bulk (median) to be reflected back into the surface layer and on into the vacuum. This bulk-bulk median interface provides the same reflective boundary as the bulk-surface slab interface does in the Bloembergen model. When considering a model describing ISOM surface science experiments, there is a physical reason why the surface and bulk regions may have different linear properties: the surface is comprised of a monolayer of atomic or molecular adsorbates sitting atop a bulk crystalline material. It is an *ad hoc* feature in this structure that the bulk and bulk median have different (linear) indices of refraction. No convincing physical argument exists as to why this should be so, however without this condition, at least in a macroscopic model, the light generated in the bulk can never be detected in the vacuum. Since this is an artificial difference, there is some question as to what the bulk-to-bulk median index ratio is. The reflection from this boundary behaves as a linear reflection. Thus, this index ratio only affects the reflected amplitude of the field generated in the bulk (median). The bulk-to-bulk median index ratio must be seen as a scaling parameter for the bulk-generated amplitude; all the angular and polarization features (for the generated waves) are unaffected by this ratio. A similar argument must be made for the thickness of the bulk median layer. Since the layer is introduced arbitrarily, the thickness is an arbitrary scaling parameter that appears primarily in the phase of the field originating from the bulk median layer. Now with light re-entering the vacuum generated from *both* the surface and the bulk (median), a comparison of the ratio of amplitudes can be made. The field amplitudes are found as they are in the Bloembergen model: all the incident, reflected, and transmitted fields are defined in the various layers and the boundary conditions are satisfied at the interfaces between layers. This operation leads to

analytical expressions for the bulk and surface generated field amplitudes. These expressions contain the incident field amplitudes, frequencies, directions, and polarizations as well as the reflected waves' directions and polarizations. The main body of this chapter explores the parameter space of these variables in a search for favourable surface-to-bulk signal ratios. This information can then be used by other researchers as a guide in doing microscopic calculations or experiments using RISOM to study surfaces.

This chapter begins with a short discussion comparing results of Ju's microscopic model with macroscopic ones. This section attempts to lend credibility to the results of macroscopic models used to describe microscopic processes. The chapter proceeds by describing macroscopic models that have been developed in the literature, particularly the model adopted and extended here; the Bloembergen model. Once the reader has been familiarized with the Bloembergen model, a discussion ensues modifying this model to include contributions from the bulk. This section develops the structure and arguments used to derive the equations for RISOM; the actual derivation of these expressions is done in a following section and is detailed in Appendix B. As it happens, it is easy *not* to assume the particular optical mixing process and derive the reflection coefficients for surface and bulk for *general optical mixing*. This is done by defining a generalized effective source wavevector and source index of refraction which are due to the induced polarization. Once the equations for general mixing are found, a simple substitution can be made to predict RISOM for specific processes such as CARS-like FWM, SHG, and SFG. The remainder of this chapter is devoted to mapping out the parameter space of the model's predictions, specifically where RISOM is likely. First only collinear phase-matched geometries are considered. Collinear phase-matching is appropriate for dispersionless materials and harmonic generation. Both polarization states,  $S$  and  $P$ , are explored. In general, it is found that improvement in the surface-to-bulk signal ratio is possible for  $S$ -polarization when the reflection angle is large; for  $P$ -waves, enhancement is done by making use of the nonlinear Brewster's angle -- the nonlinear analogue to the linear Brewster's angle. Secondly, where the linear dispersion of materials is important,

non-collinear phase-matching, is considered. Here the angular dependencies for the two polarization states for SHG (where dispersion is very important, but phase-matching is automatic) and CARS-like FWM (where dispersion is less important, but phase-matching is critical) are specifically considered. It is found that the results for enhancing the signal-to-noise ratio for collinear phase-matching are the same for the non-collinear case, however, enhancement is generally better if the surface and bulk dispersions are dissimilar. In addition, it is shown that there is an angular separation of bulk and surface signals due to dispersion; if the index of refraction is different in the surface and bulk for a given frequency, the optical process in the surface may be phase-matched, whereas the bulk may not be. Hence, waves generated in the surface and bulk travel in (slightly) different directions. Once the field geometries for RISOM are found, there is some discussion about how to use these results to guide microscopic calculations and how to experimentally search for resonant RISOM signals.

### 3.2 Microscopic Versus Macroscopic Treatments of Optical Mixing

An exact *macroscopic* treatment of general optical mixing at a dielectric interface will be presented later in this chapter. However, nonlinear processes (optical mixing) are best described by a microscopic model. The validity of using a macroscopic model has been addressed by Ju [22]. His analysis of sheet dipole model shows the classic linear Fresnel coefficients and Bloembergen's macroscopic nonlinear thin slab model yield results that exhibit the main features of a microscopic treatment. He shows that, to good approximation, a macroscopic model can furnish the linear and nonlinear reflection coefficients.

Ju performs a microscopic calculation of the optical response of a semi-infinite medium in a discrete point-dipole model. The dipoles, free to oscillate, are arranged in layers as pictured in Figure 3.2. The semi-infinite bulk is modeled as an infinite number of

dipole sheets; each sheet having the same (microscopic) polarizability tensor. In this framework, the surface SHG (SSHG) field is found by integrating the radiation field from a continuous surface sheet of nonlinear dipoles. Below the surface sheet of dipoles is positioned a semi-infinite series of (bulk) dipole sheets. Each individual induced microscopic nonlinear dipole moment constituting the nonlinear dipole sheet inside the medium is given by

$$\vec{P}^{(2)} = \gamma^{(2)}(2\omega; \omega, \omega) : [\vec{E}_{local} e^{ik_x x}][\vec{E}_{local} e^{ik_x x}] \delta(z - 0^-) \quad (3.1)$$

where  $\vec{P}^{(2)}$  is the (second order) nonlinear dipole moment,  $\vec{r} = (x, y, z)$ ,  $\gamma^{(2)}$  is the (second order) microscopic susceptibility of the sheet. The incident field is degenerate in wavevector and is incident in the  $x$ - $z$  plane.

$\vec{E}_{local}$  is the local field at the location of an individual dipole. The local field is field applied to a point-dipole due to the incident field augmented by the presence of a medium (a semi-infinite number of dipole sheets). Ju calculated this field numerically by first considering the response of a medium to an external field via an induced (linear) dipole moment. The medium is approximated by infinite sheets of point dipoles (see Figure 3.2) each responding and adding to the local field at a particular site (e.g., a lattice site) inside the medium. The induced dipole moment at a lattice site  $i$  with a (linear) dipole polarizability,  $\gamma_{dipole,i}^{(1)}$ , is

$$\vec{p}(\vec{r}_i, t) = \gamma_{dipole,i}^{(1)} \cdot \vec{E}_{local}(\vec{r}, t) = \gamma_{dipole,i}^{(1)} \cdot \left[ \vec{E}_{external} e^{i(\vec{k} \cdot \vec{r}_i)} + \sum_{i \neq j} \vec{E}_{dipole,j}(\vec{r}_i, t) \right], \quad (3.2)$$

where the local electric field is the superposition of the external field and the electric dipole field from all lattice sites in the crystal except the site itself. The local field is solved self-consistently. Once the local field due to the external field and the presence of the medium is known, the *nonlinear* response to the medium can be calculated via (3.1).

The electric dipole field of the *nonlinear* dipole moment in a continuous medium of refractive index  $n$  at frequency  $2\omega$  is found realizing that [74]

$$\vec{E}_{dipole,2\omega}(\vec{R}, \vec{r}) = \frac{1}{n^2} \vec{\nabla} \times \vec{\nabla} \times \vec{P}^{(2)}(\vec{r}) \frac{e^{i\vec{k}_{2\omega} \cdot |\vec{R} - \vec{r}|}}{|\vec{R} - \vec{r}|}.$$

Here the magnitude of the wavevector is

$$|\vec{k}_{2\omega}| = \frac{2\omega}{c} n.$$

The observed field far away from the dipoles (at location  $\vec{R}$ ) is integrated over all the contributions to the field by each dipole. With  $N$  being the surface density for each dipole sheet, this integration appears as

$$\vec{E}_{observed}(\vec{R}) = N \int d\vec{r} \vec{E}_{dipole,2\omega}(\vec{R}, \vec{r}).$$

In summary, the surface local field (as calculated numerically using linear dipole moments) and the surface nonlinear susceptibility tensor determine the nonlinear dipole moments. The dipolar interaction between the *linear dipole moments* (at the fundamental frequency in the medium) requires a self-consistent treatment in order to obtain the correct medium response to the electric dipole fields of the *nonlinear dipole moments* (oscillating at the harmonic frequency). Hence, mapping out the parameter space of polarizations and incident angle combinations would be a time-consuming task, each point in this space requiring a lengthy self-consistent calculation.

Ju's analysis produces the microscopic surface local field in the linear case. The model shows that the macroscopic results, though approximate, exhibit the important features of the microscopic field. For instance, Figure 3.5 shows the (real part of the linear) surface local field for the reflection of *P*-waves scattering off the interface (at an arbitrary angle of 45 degrees) between a vacuum and 100 dipole sheet layers (a distance equivalent to one wavelength of light used in the computation). Note that surface effects are evident. As shown by Ju, when these surface effects are averaged (over distance into the layered structure) the macroscopic result is reproduced. The good approximation of the macroscopic field to the microscopic field is illustrated in Figure 3.6. The ratio of the (real part of the) microscopic reflected field amplitude to the (real part of the) incident field amplitude is plotted as a function of angle. Here the dependence is shown about Brewster's angle. The reflectivity of *P*-waves for two wavelengths are shown: the lower curve represents a wavelength 10 times longer than the upper curve of wavelength equivalent to  $2eV$ . The index of refraction was 4.0 and the distance between layers was 0.25 nm. Note that the microscopic field has a minimum at Brewster's angle, but never reaches zero. Also in Figure 3.6, the lowest curve represents the *P*-wave macroscopic field.



The macroscopic field is given by the familiar Fresnel reflection coefficient, hence the curve is labeled 'Fresnel result'. The Fresnel result predicts, incorrectly, that the field goes to exactly zero at Brewster's angle. Despite this anomaly, the macroscopic field behaviour closely follows the microscopic local field. This illustration is representative of the many comparisons Ju's work makes between macroscopic and microscopic fields.

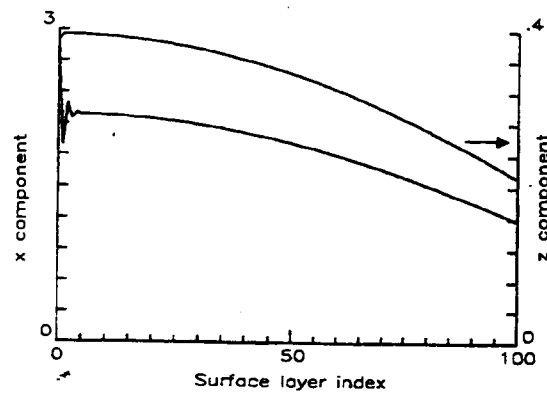


Figure 3.5: A plot of the real part of the microscopic field versus distance into a layered dipole sheet structure.

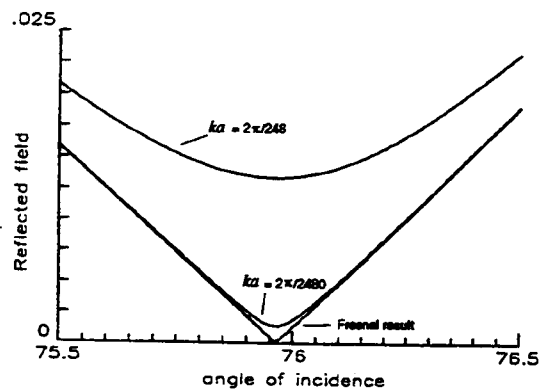


Figure 3.6: The reflectivity of a dielectric as calculated using a layered dipole sheet structure.

The conclusion drawn from Ju's work is that the familiar results derived by Maxwell's equations using macroscopic fields are consistent with, but approximate to, the results using microscopic fields when scattering off a vacuum-dielectric interface. This lends justification to using Maxwell's equations and the macroscopic fields to derive linear and nonlinear scattering off dielectric interfaces. Thus macroscopic models can be used to describe fundamentally microscopic processes (like ISOM), although the results they yield are approximate. Ju shows that macroscopic treatments by Bloembergen [23] and Sipe [75] yield very acceptable results for optical processes with ultimately microscopic origins; this is a key point for this work. Drawing from Ju, the macroscopic treatment of Bloembergen can be used to approximately describe RISOM.

### 3.3 Macroscopic Models Describing Nonlinear Processes at Surfaces

Before developing a model that predicts the features of surface *and* bulk nonlinear reflections, it is informative to review the literature. There are only two macroscopic models that are useful for the studies done here. These two treatments have some conflicting predictions that are addressed. The direction taken by this research was to choose one of these treatments and modify it to include contributions from the bulk as well as the layer at the surface. After one treatment is shown to be more appropriate for this work, a development for a RISOM model is introduced.

#### 3.3.1 Previous Work

There are two popular macroscopic models reported in the literature relevant to the work presented here. They are the treatments of Bloembergen done in the early sixties [23] and Sipe [75] published in 1987. In the exact macroscopic RISOM model presented here, the Bloembergen approach is used, hence a description of that approach is furnished. Since the treatment of Sipe is very different and some of the results do not agree with Bloembergen's, it is of interest to briefly review and contrast his approach.

In the Bloembergen development, the interfacial region is divided into three parts: the top (or cover) above, the semi-infinite bulk below, and the intermediate surface region. In this model, the cover and bulk are given only linear susceptibilities, whereas the surface layer is given both linear and nonlinear properties. There is an inhomogeneous term in the optical wave equation caused by the nonlinear polarizability radiating a wave from the surface. Using Maxwell's equations and the boundary conditions at the defined dielectric perimeters, Bloembergen derives nonlinear reflection and transmission coefficients. With these in hand, the treatment progresses by taking the limit as the depth of the surface region become small compared with the wavelengths. This is accomplished by expanding the coefficients to first order in  $d/\lambda$ , where  $d$  is the surface layer thickness. The treatment correctly predicts the linearly increasing generated field strength and the existence of a nonlinear Brewster's angle. Also predicted is the *Secant* behaviour of the reflected field magnitude on angle. According to this theory, the integral over the region of overlap of the applied fields (commonly called the interaction region) is not necessary since the limit over  $d$  is taken. Taking this limit before performing the integration leaves an expression that neglects some of the phase-matching behaviour, specifically the *Sinc* behaviour of the field strength on the phase-matching parameter. It is unimportant in Bloembergen's analysis as phase-mismatched processes are never considered.

Mizrahi and Sipe [76] took a different phenomenological approach. They treated the region that contributes to surface-generated SHG (SSHG) as an induced (continuous) nonlinear polarization sheet sitting between the vacuum at  $z = 0^+$  and a linear medium at  $z = 0^-$  (Figure 3.7). The nonlinear polarization sheet is induced by the fundamental field in the medium ( $z = 0^-$ ) with a surface nonlinear susceptibility tensor  $\gamma^{(2)}(2\omega; \omega, \omega)$ . The electromagnetic field generated by the induced dipole sheet, which is treated as a macroscopic source term, is obtained from a Green's-function formalism [75]. The reflected SSHG has two contributions, the directly generated upward-propagating wave and the downward-propagating wave reflected upward by the vacuum-medium interface at

$z = 0$  (Figure 3.7). The transmitted SSHG is just the downward-propagating wave picking up the (linear) Fresnel transmission coefficient at the interface. Their result in  $S$ -polarization is identical to that of Bloembergen (see Equations (3.17) through (3.19)). Their  $P$ -polarization results (using the definition (3.22)) for reflection and transmission, respectively, are:

$$\mathcal{E}_{r,P} = -\mathcal{E}_{s,P} \frac{(n_t^2 - n_s^2)(\cos\theta_t \sin\alpha + n_t^2 \sin\theta_t \cos\alpha)}{n_t^2(\cos\theta_r + \cos\theta_t)}$$

and

$$\mathcal{E}_{t,P} = -\mathcal{E}_{s,P} \frac{(n_t^2 - n_s^2)(\cos\theta_r \sin\alpha - n_t \sin\theta_t \cos\alpha)}{n_t^2(\cos\theta_r + \cos\theta_t)}.$$

When compared to Bloembergen's results (refer to Equations (3.20) through (3.22)), the contribution from  $z$ -component of nonlinear polarization (e.g., the term involving  $\mathcal{E}_{s,P} n_t \sin\theta_t \cos\alpha$ ) differs by a factor of  $n_t^2$ , the index of refraction in the medium at the frequency of the second harmonic. This discrepancy is the result of bringing the nonlinear polarization source *outside* the linearly responding medium. Recall that the tangential component of the electric field and the normal component of the displacement field ( $\bar{D} \equiv n^2 \bar{E}$ ) are continuous across the interface. The model considered by Mizrahi and Sipe is reasonable only when the nonlinearity arises mainly from molecules adsorbed on a surface and the local field effects of the molecules can be neglected. This is justified if surface coverage of adsorbed molecules is so low that the effective dielectric constant of the adsorbing layer is very nearly one. Bloembergen's model does not have this assumption, hence it was chosen as the starting place for a RISOM theory. The next section discusses Bloembergen's model in detail in preparation for the development of the RISOM theory developed here.

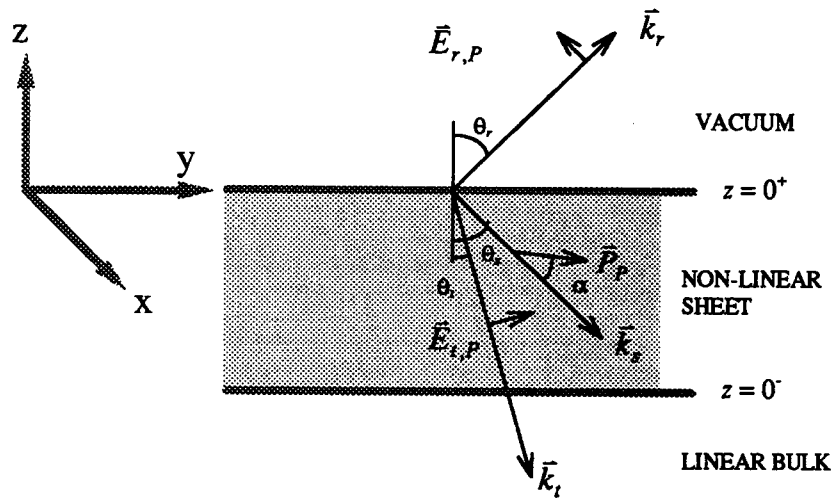


Figure 3.7 A nonlinear polarization sheet at  $z = 0^+$  using in the Mizrahi-Sipe Model for  $P$ -polarization.

### 3.3.2 Bloembergen's Model in Detail

Bloembergen and co-workers determined the theoretical equations that govern optical mixing, particularly SHG, in reflection geometry [23], [24], [25]. In his first paper, Bloembergen makes the assumption that 'mixed waves' (waves generated due to a nonlinear susceptibility) emanate or 'reflect' from the interface. These waves are produced by a source wave; it is created by a nonlinear polarization induced by all incident fields. With this, he presents the solutions to Maxwell's equations in nonlinear dielectrics. These are the solutions to Maxwell's equations which satisfy the boundary conditions at a plane interface between a linear and nonlinear medium. Generalizations to the well-known laws of (linear) reflection and refraction give the direction of propagation of the mixed waves generated in a thin region near the interface. He shows that these waves' intensity and polarization conditions are described by generalized (nonlinear) Fresnel coefficients.

This *macroscopic* treatment begins with a statement concerning mixed waves emanating from an interface. Bloembergen's development begins with a description of the physics that governs the generation of (second) harmonic waves. This description is later generalized to mixed waves, that is, all nonlinear processes. The induced polarization oscillates at the mixing frequency. The oscillating polarization will, in turn, radiate energy in the form of a traveling wave of the same frequency. The nonlinear source term for TWM is given by

$$P^{(2)}(\omega_3) = \chi^{(2)}(\omega_3; \omega_1, \omega_2): \bar{E}_1 e^{i(\bar{k}_1 \cdot \bar{r} - \omega_1 t)} \bar{E}_2 e^{i(\bar{k}_2 \cdot \bar{r} - \omega_2 t)}. \quad (3.3)$$

For general  $n$ -wave mixing, that is a process with  $n-1$  incident waves that mix to produce an  $n$ -th wave, the nonlinear source term is

$$P^{(n-1)}(\omega_n) = \chi^{(n-1)}(\omega_n; \omega_1, \omega_2, \dots, \omega_{n-1}): \bar{E}_1 e^{i(\bar{k}_1 \cdot \bar{r} - \omega_1 t)} \bar{E}_2 e^{i(\bar{k}_2 \cdot \bar{r} - \omega_2 t)} \dots \bar{E}_{n-1} e^{i(\bar{k}_{n-1} \cdot \bar{r} - \omega_{n-1} t)}. \quad (3.4)$$

It is useful to define an *effective wavevector* for the source term,  $\bar{k}_s$ , as the vector sum of the wavevectors for all incident waves such that

$$\bar{k}_s \equiv \sum_{i=1}^{n-1} \bar{k}_i, \quad (3.5)$$

and an *effective index of refraction* associated with the source wave such that

$$n_s \equiv \frac{c}{\omega_n} |\vec{k}_s| = \frac{c}{\omega_n} \left| \sum_{i=1}^{n-1} \vec{k}_i \right|.$$

Then, using (3.5), expression (3.4) can be concisely written as

$$P^{(n-1)}(\omega_n) = \chi^{(n-1)}(\omega_n; \omega_1, \omega_2, \dots, \omega_{n-1}): \vec{E}_1 \vec{E}_2 \dots \vec{E}_{n-1} e^{i(\vec{k}_s \cdot \vec{r} - \omega_n t)} \quad (3.6)$$

Note that the convention for  $\omega_n$  is the *sum* of frequencies of *annihilated* photons. If there are photons emitted in the process, the frequency of those photons are treated as if they were negative. Hence, in a CARS process where  $\omega_1$  and  $\omega_3$  are annihilated, but  $\omega_2$  (as well as  $\omega_4$ ) is emitted,  $\omega_n = \omega_4 = \omega_1 + (-\omega_2) + \omega_3$ . Note that the convention also affects the wavevectors of any emitted waves. (For the previous example, the wavevector for the wave at  $\omega_4$  is  $\vec{k}_4 = \vec{k}_1 - \vec{k}_2 + \vec{k}_3$ .)

The nonlinear source term arising from the induced polarizability was introduced by Bloembergen and co-workers [77], who showed that the effective nonlinear source term can readily be incorporated into Maxwell's equations for a nonlinear medium,

$$\vec{\nabla} \times \vec{E} = -\frac{1}{c} \partial_t \vec{B}, \quad (3.7)$$

and

$$\vec{\nabla} \times \vec{H} = \frac{1}{c} \partial_t \vec{D} + \frac{4\pi}{c} \partial_t \vec{P}, \quad (3.8)$$

understanding that

$$\bar{P} = \bar{P}^{(1)} + \bar{P}^{(2)} + \bar{P}^{(3)} + \dots + \bar{P}^{(n)} + \dots, \quad (3.9)$$

Assuming that the permittivity,  $\epsilon$ , is a scalar and assuming a non-magnetic material such that the permeability,  $\mu$ , is unity, the waves at the generated frequency obey the wave equation

$$\bar{\nabla} \times \bar{\nabla} \times \bar{D}(\bar{r}, t) + \left( \frac{n(\omega_n)}{c} \right)^2 \partial_t^2 \bar{D}(\bar{r}, t) = 4\pi \bar{\nabla} \times \bar{\nabla} \times \bar{P}(\bar{r}, t). \quad (3.10)$$

Consistent with dielectrics, it is assumed no currents are induced. Also, it is assumed that the response of the system to the applied frequencies is small, hence  $\partial_t n = 0$ . Each of the fields can be (represented as harmonic fields and) transformed from  $\bar{D}(\bar{r}, t)$  into  $\bar{D}(\bar{r}, \omega)$  by defining the Fourier transforms

$$\begin{aligned} \bar{D}(\bar{r}, t) &= \left( \frac{1}{2\pi} \right)^{\frac{1}{2}} \int_{-\infty}^{\infty} d\omega \bar{D}(\bar{r}, \omega) e^{-i\omega t} \\ \bar{P}(\bar{r}, t) &= \left( \frac{1}{2\pi} \right)^{\frac{1}{2}} \int_{-\infty}^{\infty} d\omega \bar{P}(\bar{r}, \omega) e^{-i\omega t}, \end{aligned}$$

which reduce the wave equation to the Helmholtz equation with an inhomogeneous term:



$$\bar{\nabla} \times \bar{\nabla} \times \bar{D}(\bar{r}, \omega) + k^2(\omega_n) \bar{D}(\bar{r}, \omega) = 4\pi \bar{\nabla} \times \bar{\nabla} \times \bar{P}(\bar{r}, \omega). \quad (3.11)$$

A complete derivation of the nonlinear Helmholtz equation can be found in Appendix B.

It should be noted that this is the usual linear Helmholtz equation augmented by a source term on the right-hand side of the equation. The general solution of (3.10) consists of the homogeneous equation plus one particular solution of the inhomogeneous equation. A complete argument used to find a general solution is found in Appendix B. The general solution to the inhomogeneous (induced) electric fields for the two polarizations, *S* and *P* (see the discussion of Figure 3.9 and Figure 3.10 below), may be written

$$\bar{E}_{INHOMO,S}(\bar{r}, t) = \bar{\epsilon}_S \mathcal{E}_{s,S} e^{i(\bar{k}_s \cdot \bar{r} - \omega_s t)} \quad (3.12)$$

and

$$\bar{E}_{INHOMO,P}(\bar{r}, t) = \bar{\epsilon}_P \mathcal{E}_{s,P} e^{i(\bar{k}_s \cdot \bar{r} - \omega_s t)}. \quad (3.13)$$

The induced field amplitudes for the two polarizations are defined by (refer to Appendix B)

$$\bar{\epsilon}_S \mathcal{E}_{s,S} \equiv \frac{-4\pi}{n_t^2 - n_s^2} \left( \bar{P}_S^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}) - \frac{\bar{k}_s (\bar{k}_s \cdot \bar{P}_S^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}))}{|\bar{k}_s|^2} \right) \quad (3.14)$$

and

$$\bar{\epsilon}_P \mathcal{E}_{s,P} \equiv \frac{-4\pi}{n_t^2 - n_s^2} \left( \bar{P}_P^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}) - \frac{\bar{k}_s (\bar{k}_s \cdot \bar{P}_P^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}))}{|\bar{k}_s|^2} \right). \quad (3.15)$$

As shown in Appendix B, the total *observed* field in the vacuum due to an arbitrary structure of nonlinear media is the sum of a homogeneous solution and the inhomogeneous solutions to the wave equation:

$$\vec{E}(\vec{r}, t) = \vec{E}_i \mathcal{E}_i(\vec{r}, t) + \vec{E}_s \mathcal{E}_s(\vec{r}, t) \quad (3.16)$$

In (3.16) the second term is given by (3.14) and (3.15), but the first term containing  $\mathcal{E}_i$  must be found by matching the boundary conditions of the layered structure.

Equations (3.12) through (3.15) represent the nonlinear source terms for  $n$ -mixing in a nonlinear medium. With these general solutions, specific cases which include boundaries between the vacuum and a nonlinear medium are solved. As will be seen, the nonlinear Fresnel coefficients are found by considering this very geometry. It is pictured in Figure 3.8. The wavevectors ( $k_i$ ,  $k_r$ , and  $k_s$ ), polarization vectors ( $\epsilon_i$ ,  $\epsilon_r$ , and  $\epsilon_s$ ), and magnitudes of the fields ( $\mathcal{E}_i$  and  $\mathcal{E}_r$ ), are determined from these boundary conditions. It turns out that the nonlinear polarization radiates back into the vacuum in one particular direction. Likewise, because of the linear index of refraction, the polarization radiates into a certain direction into the bulk medium. The problem is analogous to the linear case of reflection and refraction at a dielectric interface. The difference is that the role of the incident wave has been replaced by the 'inhomogeneous wave' with an amplitude proportional to the nonlinear polarization,  $P$ . Note that the nonlinear Fresnel coefficients can be defined as the ratio of the nonlinear reflected amplitude to the polarization. This solidifies the analogue between the linear and nonlinear process; the linear Fresnel coefficients are defined as the ratio of the reflected to incident field amplitudes.

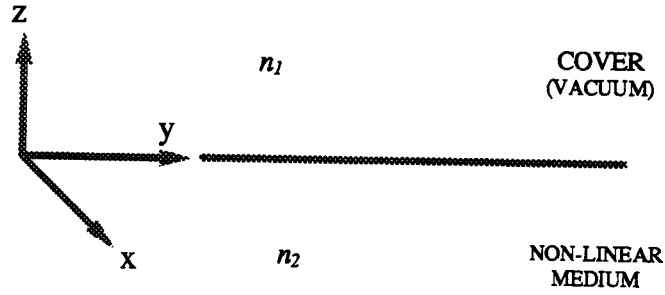


Figure 3.8: Interface between the vacuum and a nonlinear medium.

Considering the interface pictured in Figure 3.8, the tangential component of  $E$  and  $H$  should be continuous everywhere on the boundary at all times (recall infinite harmonic plane waves were assumed). This requires that the field components for each frequency (all incident and generated) be separately continuous across the boundary. For perfect phase-matching, satisfying this condition means that the  $y$ -components of the wavevectors for the generated reflected field,  $k_r$ , the transmitted field,  $k_t$ , and the effective source wavevector (recall (3.5)),  $k_s$ , are all equal:

$$(\vec{k}_r)_y = (\vec{k}_t)_y = (\vec{k}_s)_y$$

or

$$|\vec{k}_r| \sin \theta_r = |\vec{k}_t| \sin \theta_t = |\vec{k}_s| \sin \theta_s.$$

The reader may wish to refer to Figure 3.8 or Figure 3.9. Hence there are unique

directions for the reflected and transmitted generated waves. It should be pointed out that in general, the indices are *not* equal:

$$n_r \neq n_t \neq n_s.$$

Bloembergen implicitly assumed that only perfect phase-matching cases were of interest. When the process is not phase-matched, there is still a unique direction which is found by minimizing the phase difference [78]. That is, the intensity is reduced by the phase-mismatch as characterized by the difference in phase between the source and transmitted waves,  $\Phi = i(\Delta\vec{k} \cdot \vec{r})$  with  $\Delta\vec{k} \equiv \vec{k}_t - \vec{k}_s$ . Recall that  $\vec{k}_s$  depends only on fixed constants and the incident field parameters (which can be considered fixed for argument's sake). The transmitted wavevector,  $\vec{k}_t$ , depends on the frequency, the index of refraction of the medium at that frequency (which are fixed) as well as the direction of the wave. Hence, the only free parameter available to minimize the phase difference is the direction of the transmitted wave. This is how a unique direction for this wave arises. Notice, from (3.12) and (3.13) and the boundary conditions, nonlinear Fresnel coefficients can be defined. Bloembergen derives these expressions and comments on general laws of nonlinear reflection and refraction for a non-dispersive, phase-matched process.

The general laws of reflection and refraction can be determined by considering a simple vacuum-nonlinear medium interface (Figure 3.8). From this structure, Bloembergen derives the nonlinear Fresnel coefficients for an isotropic medium. Considering an isotropic medium is important for the work presented here. Essentially this assumption means that the nonlinear susceptibility is taken to be a unit tensor. Later this will be important for the comparison of surface and bulk contributions to the generated signal without the need to consider the (possibly different) symmetry of surface and bulk. This is a worst case scenario for RISOM. If, in real experiments, there *are*

convenient symmetries that enhance the surface over the bulk signals, they would certainly be used. For argument's sake, no symmetry assumptions will be made.

Bloembergen derives the field amplitudes for  $S$ -waves using continuity of the tangential components of  $E$  and  $H$  across  $z = 0$ . Combining these boundary conditions with the solution to the inhomogeneous problem (from (3.12) through (3.15) and results from Appendix B) yields the amplitudes of the reflected and transmitted fields. The coordinate system is shown in Figure 3.9.

With reflected wavevectors (with directions described by angles  $\theta$  with respect to the  $z$ -axis) and field amplitudes labeled with subscripts  $r$ , transmitted labeled with subscripts  $t$ , and source waves labeled with  $s$ , the field amplitudes for  $S$ -wave are given by

$$\mathcal{E}_{r,S} = \mathcal{E}_{t,S} - \frac{4\pi P_S}{n_t^2 - n_s^2} = \mathcal{E}_{t,S} - \mathcal{E}_{s,S} = \mathcal{E}_{s,S} \frac{n_t \cos \theta_t - n_s \cos \theta_s}{n_t \cos \theta_t + n_r \cos \theta_r} \quad (3.17)$$

$$\mathcal{E}_{t,S} = \mathcal{E}_{r,S} + \mathcal{E}_{s,S} \left( e^{i(\bar{k}_s - \bar{k}_t) \cdot \bar{r}} - 1 \right) \quad (3.18)$$

with

$$\mathcal{E}_{s,S} \equiv \frac{4\pi P_S}{n_t^2 - n_s^2}. \quad (3.19)$$

Here the capital subscript ' $S$ ' refers to perpendicularly ( $S$ ) polarized waves.

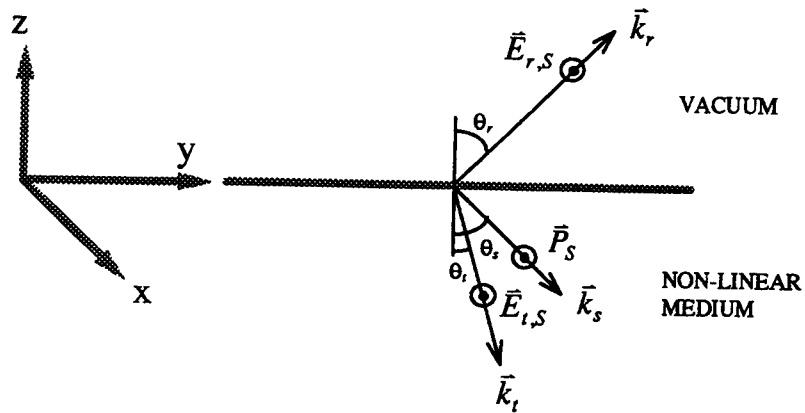


Figure 3.9: Generated waves at the boundary of a nonlinear medium, polarized with the electric field vector normal to the plane of reflection.

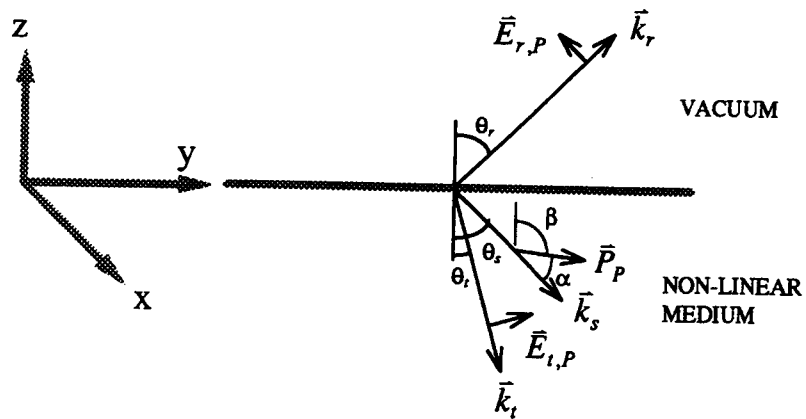


Figure 3.10: Generated waves at the boundary of a nonlinear medium, polarized with the electric field vector in the plane of reflection.

The  $P$ -waves are those generated by the  $P$ -component of the induced polarization, that is, the  $y$  and  $z$ -components in the coordinate system defined in Figure 3.10. It is advantageous to describe the component of the nonlinear polarization in the plane of reflection by its magnitude and angle  $\alpha$  between its direction and the direction of propagation of the source wave, described by  $\vec{k}_s$ . Again, the amplitudes are derived considering the continuity of the tangential components of  $E$  and  $H$  (from (3.12) and (3.13)) at  $z = 0$ . Thus the field amplitudes for  $P$ -waves are:

$$\mathcal{E}_{r,P} = -\mathcal{E}_{s,P} \frac{(n_t^2 - n_s^2) \sin \theta_s \sin^2 \theta_t \sin(\alpha + \theta_r + \theta_s)}{n_r^2 \sin \theta_r \sin(\theta_t + \theta_s) \sin(\theta_t + \theta_r) \cos(\theta_t - \theta_r)}, \quad (3.20)$$

$$\mathcal{E}_{t,P} = -\mathcal{E}_{s,P} (n_t^2 - n_s^2) \left( \frac{\frac{n_r^2 \sin \theta_r}{n_t^2 \sin \theta_t} + \frac{\sin \alpha \sin \theta_s \cos \theta_s}{n_t^2 \sin(\theta_t + \theta_s)} + \frac{\cos \alpha \sin(\theta_t - \theta_s)}{n_t^2} e^{i(\vec{k}_s - \vec{k}_t) \cdot \vec{r}} + \frac{\sin \alpha \cos(\theta_t - \theta_s)}{(n_t^2 - n_s^2)^2} (e^{i(\vec{k}_s - \vec{k}_t) \cdot \vec{r}} - 1) \right), \quad (3.21)$$

$$\mathcal{E}_{s,P} \equiv \frac{4\pi P_P}{n_t^2 - n_s^2}. \quad (3.22)$$

Note that in the case of  $S$ -polarization, there is no advantage to considering

$\bar{k}_s \cdot \bar{P}_P^{(n)} \neq 0$  since the dipole sheet radiation strength changes only with the *cosine of the polar angle*  $\theta$ . This is why no angle  $\alpha$  is considered for the *S*-polarization case (Figure 3.9). In contrast, for *P*-polarization,  $\bar{k}_s \cdot \bar{P}_P^{(n)} \neq 0$  is important to consider. When the nonlinear dipole moments are aligned in the *P*-plane, it is possible that the applied fields are in the same direction as the dipole moments. By analogy with the linear case, this angle is called the nonlinear Brewster's angle (NLBA). The NLBA is found by examining the last factor in the numerator of (3.20). Note that  $\theta_r = \pi - \alpha - \theta$ , is the angle for reflected wave and when  $\alpha = 0$  (when  $\bar{P}_P^{(n)}$  is parallel to  $\bar{k}_s$ ), this factor in causes the amplitude to vanish. The physical interpretation of the NLBA is that the nonlinear polarization cannot radiate in the direction it oscillates. In Appendix B, there is a more detailed discussion as to the physical meaning of the NLBA.

Equations (3.17), (3.18), (3.20), and (3.21) give the generated wave field amplitudes when fields are applied to an interface between a linear and a semi-infinite nonlinear material. Bloembergen was interested in obtaining expressions for these amplitudes when the nonlinear material is a thin slab rather than a semi-infinite dielectric. He proposed the structure shown in Figure 3.3. It consists of a thin slab of material with both linear and nonlinear properties atop a semi-infinite bulk *with only linear* properties. In this approach, Bloembergen assumed optical mixing occurred within the thin slab of nonlinearly active material.

The layered structure with its coordinate system is defined in Figure 3.11. At the cover-surface layer interface,  $z = 0$ , and at the surface-bulk interface  $z = -d_2$ . Thus, the surface layer thickness is  $d_2$ . Each region is has a unique linear index of refraction indicated by  $n_1$  for the cover,  $n_2$  for the surface, and  $n_3$  for the bulk region. The fields are defined in each region by considering the *S* and *P* polarizations (see Figure 3.9 and Figure 3.10) and by the wavevectors. The fields have amplitudes  $\mathcal{E}$  and are defined with zero phase at the top of each interface. Hence, after propagating through a layer of thickness  $d$ , they will acquire a phase factor  $e^{i\bar{k} \cdot d\hat{z}}$ .



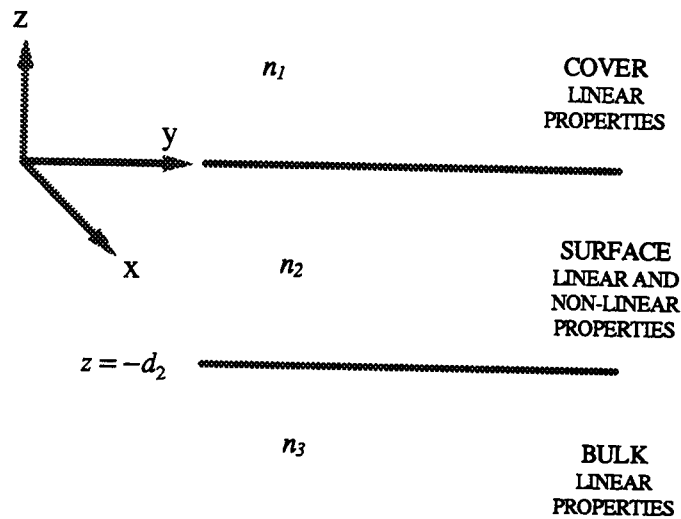


Figure 3.11: Bloembergen's layered structure with its coordinate system.

Recall that the wavevectors in each region are dependent on the linear frequency-dependent indices of refraction. The wavevectors are defined in Figure 3.12. (This may be compared with Figure 3.3.) The field in each region propagates with an associated wavevector is given identifying subscripts. The wavevectors and fields are identified with the same subscripts. These subscripts are defined by one letter describing the frequency of the wave, followed by a second letter identifying the wave (I = incident, r = reflected, t = generated, s = source, u = transmitted), and a number indicating the layer. The layer numbers are: cover layer (vacuum) = 1, thin surface slab = 2, semi-infinite bulk = 3.

For the  $n$ -I incident waves, the number indicates the layer in which the wave is located, as shown at the far left of Figure 3.12. For instance, the applied wave 'I' with frequency  $\omega_I$ , reflected from the bulk-surface interface and traveling in the surface layer, has associated subscripts 'I,r2'. If the process being considered was SFG, there would be two incident waves with associated subscripts 'I,I1' and '2,I1'. The reflection of these

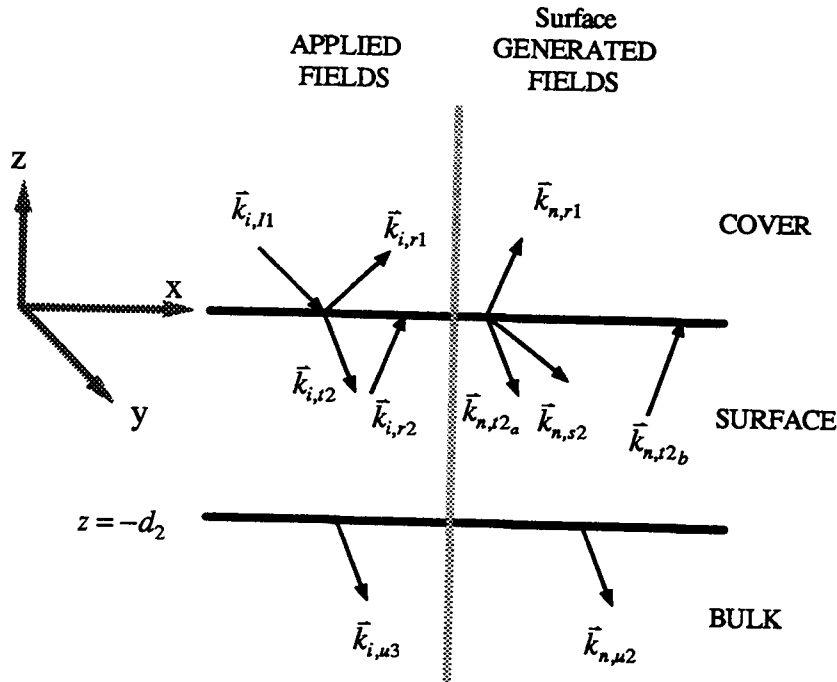


Figure 3.12: Defining the wavevectors for the Bloembergen model.

waves in the cover layer are labeled ' $l,r1$ ' and ' $2,r1$ ', whereas the transmitted waves have labels for the surface layer are ' $l,t2$ ' and ' $2,t2$ ', and so forth. Explicitly,  $\vec{k}_{i,1}$  is the wavevector of the  $i$ -th incident wave with frequency  $\omega_i$ . It defines one of the  $n-1$  waves incident from the cover heading to the slab interface.  $\mathcal{E}_{i,1}$  is the amplitude of the field of this wave.  $\vec{k}_{i,r1}$  is the wavevector of the  $i$ -th wave reflected from the interfaces below. This wave returns to the cover (vacuum);  $\mathcal{E}_{i,r1}$  is the amplitude of this wave.  $\vec{k}_{i,t2}$  is the wavevector of the  $i$ -th wave transmitted through the surface layer down towards the semi-infinite bulk;  $\mathcal{E}_{i,t2}$  is its field amplitude.  $\vec{k}_{i,r2}$  is the wavevector of the  $i$ -th wave which reflects from the surface-bulk interface and travels up towards the cover;  $\mathcal{E}_{i,r2}$  is the electric field amplitude of this wave.  $\vec{k}_{i,\mu3}$  is the wavevector of the  $i$ -th wave that is

transmitted into the semi-infinite bulk.  $\mathcal{E}_{i,u3}$  is the amplitude of the field of this wave. Each wavevector is described by spherical coordinates,  $\vec{k}_\lambda = (|\vec{k}_\lambda|, \theta_\lambda, \phi_\lambda)$ , with  $\lambda$  equal to any of the subscripts defined for the wavevectors and amplitudes above.

For the generated waves, the layer identification number indicates *the layer number the wave is generated in*. (In Bloembergen's derivation the only layer that waves are generated in is layer 2 -- the surface layer.) For example, the generated wave parameters (which has frequency index  $n$ ) propagating into the semi-infinite bulk has subscripts ' $n,u2$ '. Specifically,  $\vec{k}_{n,r1}$  is the wavevector of the wave generated from the surface region. It is the wave reflected from the interfaces below and now propagates in the vacuum.  $\mathcal{E}_{n,r1}$  is the field amplitude of this wave.  $\vec{k}_{n,t2_a}$  is the wavevector of the wave generated in surface region.  $\mathcal{E}_{n,t2_a}$  is the amplitude of the field of this downwards-traveling wave. The generated wave described by  $\vec{k}_{n,t2_a}$  is reflected from the surface-bulk interface. This wave reflected from the surface-bulk interface, propagates in the surface slab; it has  $\vec{k}_{n,t2_b}$  as its wavevector;  $\mathcal{E}_{n,t2_b}$  is its amplitude.  $\vec{k}_{n,s2}$  is the wavevector of the 'source' wave. It is the effective wavevector as defined in equation (3.5). The amplitude of this wave,  $\mathcal{E}_{n,s2}$ , is given by equations (3.12) through (3.15). Finally, the wave transmitted down into the semi-infinite bulk is given the wavevector  $\vec{k}_{n,u2}$  and amplitude  $\mathcal{E}_{n,u2}$ . Again, each wavevector is described by spherical coordinates:  $\vec{k}_\lambda = (|\vec{k}_\lambda|, \theta_\lambda, \phi_\lambda)$ , with  $\lambda$  equal to any of the subscripts defined for the wavevectors and amplitudes above.

Notice that only one reflection off of each interface is included. When the reflection angle is large (near grazing) this model is suspect as the slab will act as a leaky waveguide with many reflections at each interface before the light 'escapes' from the confines of the slab. This model is also limited to 'weak coupling' theory. That is the incident waves are not attenuated by the generation of a new, inhomogeneous, wave.

Upon imposing the boundary conditions on the fields, realizing the assumptions made, four equations for each polarization arise. Bloembergen's solutions to this set of equations

for generated  $S$ - and  $P$ -polarization reflected field amplitudes from the surface slab,  $\mathcal{E}_{nr1}$ , (dropping the  $n$  subscript) are:

$$\mathcal{E}_{r1,S} = -\mathcal{E}_{s2,S} D_S^{-1} \begin{pmatrix} (n_{s2} \sin \theta_{s2} - n_{u2} \sin \theta_{u2})(\cos \phi_{i2} - \cos \phi_{s2}) \\ + i \frac{n_{u2} \sin \theta_{u2}}{n_{i2} \sin \theta_{i2}} (n_{i2} \cos \theta_{i2} \sin \phi_{s2} - n_{s2} \cos \theta_{s2} \sin \phi_{i2}) \\ + i (n_{i2} \cos \theta_{i2} \sin \phi_{i2} - n_{s2} \cos \theta_{s2} \sin \phi_{s2}) \end{pmatrix} \text{ and} \quad (3.23)$$

$$\begin{aligned} \mathcal{E}_{r1,P} = & -\mathcal{E}_{s2,P} D_P^{-1} \sin \alpha \begin{pmatrix} (n_{u2} \sin \theta_{s2} - n_{s2} \sin \theta_{u2})(\cos \phi_{i2} - \cos \phi_{s2}) \\ + i \frac{n_{u2}}{n_{i2}} (n_{s2} \cos \theta_{u2} \sin \phi_{i2} - n_{i2} \cos \theta_{s2} \sin \phi_{s2}) \\ + i \sin \theta_{u2} \sec \theta_{i2} (n_{s2} \cos \theta_{i2} \sin \phi_{s2} - n_{i2} \cos \theta_{s2} \sin \phi_{i2}) \end{pmatrix} \\ & - \mathcal{E}_{s2,P} D_P^{-1} \frac{1}{n_{i2}} \cos \alpha \begin{pmatrix} n_{u2} \sin \theta_{s2} (\cos \phi_{i2} - \cos \phi_{s2}) \\ - i \sin \theta_{s2} \sec \theta_{i2} (n_{u2} \cos \theta_{i2} \sin \phi_{s2} + n_{i2} \cos \theta_{u2} \sin \phi_{i2}) \end{pmatrix} \end{aligned} \quad (3.24)$$

with

$$\begin{aligned} D_S \equiv & \cos \phi_{i2} (n_{u2} \sin \theta_{u2} + n_{r2} \sin \theta_{r2}) \\ & - i \sin \phi_{i2} \left( \frac{n_{r2} n_{u2}}{n_{i2}} \cos \theta_{r2} \cos \theta_{u2} \sec \theta_{i2} + n_{i2} \cos \theta_{i2} \right), \end{aligned} \quad (3.25)$$

$$\begin{aligned} D_P \equiv & \cos \phi_{i2} (n_{r2} \sin \theta_{u2} + n_{u2} \sin \theta_{r2}) \\ & - i \frac{1}{n_{i2}} \sin \phi_{i2} \sec \theta_{i2} (n_{i2} \cos \theta_{r2} + n_{r2} n_{u2} \cos^2 \theta_{i2}), \end{aligned} \quad (3.26)$$

$$\phi_{s2} \equiv n_{s2} \frac{\omega}{c} d_2 \cos \theta_{s2},$$

and

$$\phi_{t2} \equiv n_{t2} \frac{\omega}{c} d_2 \cos \theta_{t2}.$$

Several observations that can be made from (3.23) through (3.26). For instance, when there is no longer any discontinuity between the surface and bulk layers (when  $n_s = n_u$ ), the reflection vanishes. Also, the surface-to-bulk index ratio is only scaling factor in the amplitude of the generated field as measured in the vacuum. Also, but not readily apparent from (3.23) through (3.25), is the famous  $\sec \theta_t$  behaviour often quoted in the literature. It can be seen by realizing that the denominators ((3.25) and (3.26)) of both (3.23) and (3.24) contain a  $\cos \theta_s$ . Thus, it is expected that the surface reflected wave should increase in intensity as the angle increases to near grazing angles.

For *P*-waves, note the occurrence of the NLBA for perfect phase-matching ( $\vec{k}_{t2} = \vec{k}_{s2}$ ). For perfect phase-matching  $\theta_{s2} = \theta_{t2} = (\pi - \alpha)/2$  and when the source wave and polarization are parallel ( $\alpha = 0$ ), the first term in (3.24) and the term involving  $\cos \phi_{s2} - \cos \phi_{t2}$  immediately vanish. The remaining term vanishes only when the limit of perfect phase-matching is taken -- the denominator tends toward infinity more rapidly than the numerator. Under these conditions there is no reflected wave.

Bloembergen takes equations (3.23) through (3.26), expands them in  $kd_2$ , and approximates them to first order in  $kd_2$  (the thickness is much smaller than the wavelength). He shows that the reflected and transmitted intensities are equal and that

they are proportional to the square of the thickness of the layer. This is consistent with observed behaviour.

### 3.3.3 Results and Limitations of Bloembergen's Model

Exact macroscopic treatments by Bloembergen of general optical mixing at a dielectric interface yields many accurate results. However, the integral over the region of overlap of the applied fields (commonly called the interaction region) is not necessary since the limit of small  $d$  is taken. Taking this limit before performing the integration leaves an expression that neglects some of the phase-matching behaviour, specifically the *Sinc* behaviour of the field strength on the phase-matching parameter. It is unimportant in Bloembergen's analysis as phase-mismatched processes are never considered. Furthermore, Bloembergen's result neglects terms that are second order in the phase-mismatch. Although it only effects the transmitted wave amplitudes (and unimportant for discussions regarding the reflected waves), it should be pointed out that the results for Bloembergen's transmitted waves ([23], equations (6.8) and (6.18)) neglect a term of the form

$$-i \frac{n_{r1} \cos \theta_{r1}}{n_{t2} \cos \theta_{t2}} (n_{t2} \cos \theta_{t2} \sin \phi_{s2} - n_{s2} \cos \theta_{s2} \sin \phi_{t2})$$

which clearly vanishes for perfect phase-matching ( $\vec{k}_{t2} = \vec{k}_{s2}$ ); it is also the only one that is second order in the phase-mismatch. When the phase-mismatch is large, this term cannot be neglected.

From the perspective of RISOM, the largest limitation to current macroscopic treatments of OM at an interface (either Bloembergen's or Sipe's) is that they only predict the surface-generated waves. There is no information concerning the bulk-generated waves. There is no apparent reason that the surface-generated signal should dominate

over the bulk signal. Yet, the search for favourable RISOM conditions relies on maximizing the surface-to-bulk signal ratio, hence it is critical that information about *both* the bulk and surface signals be available.

### 3.4 Developing a Macroscopic Model for RISOM Prediction

The structure, shown in Figure 3.4, was adopted in order to develop a RISOM model. It is carefully pictured with its coordinate system in Figure 3.13 below.

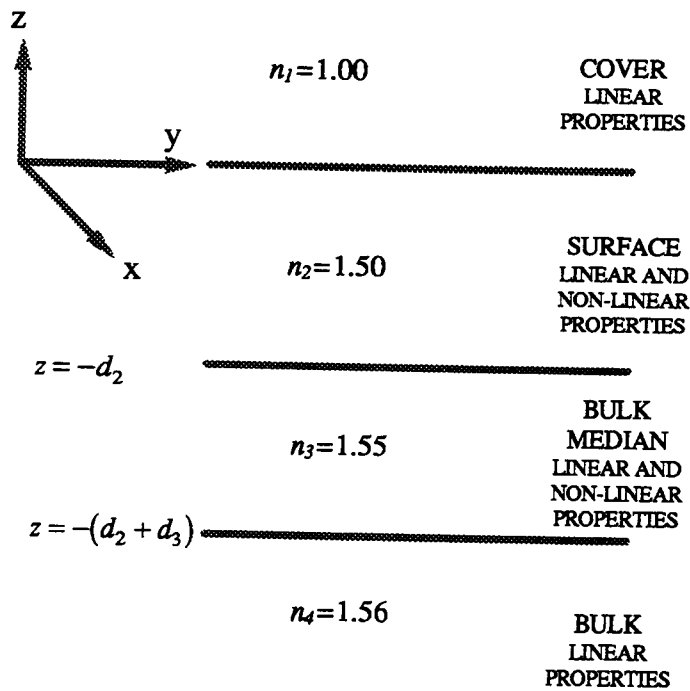


Figure 3.13: A four layer structure for developing a RISOM model.

The new model breaks the interfacial region into *four* distinct regions: a cover, a surface region, a bulk median, and the semi-infinite bulk. The additional layer is an extension to Bloembergen's model, and is added between the surface slab and the bulk. This is referred to as the *bulk median*. This layer has the nonlinear properties of the bulk

which are, in general, different from the surface. The linear index of refraction of the bulk median, however, is made *different* than the semi-infinite bulk. This new interface establishes a mechanism for the light generated in the bulk (median) to be reflected back into the surface layer and on into the cover (e.g., vacuum). This bulk-bulk median interface provides the same reflective boundary as the bulk-surface slab interface does in the Bloembergen model.

When describing surface-specific experiments on real surfaces (including ISOM experiments), there are physical reasons why the surface and bulk regions should have different linear and nonlinear properties. A surface includes an interstitial region comprised typically one or two atomic layers between the bulk and the cover. Within this interstitial region there are several influences that may cause the linear and nonlinear properties of this region to be distinct from the bulk crystal. Relaxation of the unit cell parameters may occur throughout this region as the bulk stoichiometry re-organizes to fit the stoichiometry of the cover (the cover is usually considered to be a vacuum, however the possibility of it being another linear dielectric is left open). Impurities are likely to be found in higher concentrations in the surface region; they are left there during sample preparation and handling. Because of these idiosyncrasies between the surface and bulk stoichiometry, the surface region is expected to have a rather different electric and magnetic (linear and nonlinear) properties than the bulk. It can be expected that the linear index of the bulk changes gradually over several atomic layers below the surface. The bulk median simulates this gradation.

The addition of an interface consisting of different linear indices of refraction between the bulk and bulk median plays the same role as Bloembergen's interface between the surface and bulk. This new interface provides a mechanism by which the light generated in the bulk median can return to the vacuum. Now the surface-to-bulk amplitude ratio for



the waves generated in both regions can be analyzed for different parameters. Similar to Bloembergen's surface layer, that the bulk-to-bulk median index ratio and the thickness of the bulk median layer are considered a scaling parameters to the bulk-generated amplitude.

### 3.4.1 Deriving the Behaviour of the Generated Fields

The same method to derive the generated fields for the surface and bulk median layers as was used in the Bloembergen model. Reviewing equation (3.15), the reflected nonlinear fields are found knowing the inhomogeneous source wave solutions (equations (3.13) and (3.14)) and the boundary conditions of the layered structure of the model. A complete derivation of the reflected nonlinear fields are given in Appendix B. A conceptual diagram of this structure is shown in Figure 3.4. A working diagram of the layered structure with its coordinate system is defined in Figure 3.13. The boundary conditions for each boundary indicated in Figure 3.13 are used to obtain the expressions for the fields generated in the surface and bulk median layers. Note that Figure 3.13 is divided into three sections left to right. This is done to emphasize that there are three independent sets of fields found in all parts of the structure (vertically): the applied fields, the fields generated in the surface region, and the fields generated in the bulk median region. Indicated between the layers are the wavevectors for the defined fields; each field is defined with zero relative phase at the upper surface for each layer. Only one reflection from the bottom layer of each layer is included; multiple reflections are ignored. Matching the fields at the boundaries of the regions of the new structure yields expressions for the reflected generated electric fields (entering the cover region) from the surface and bulk median layers.

At the cover-surface layer interface,  $z = 0$ , at the surface-bulk median interface  $z = -d_2$ , and at the semi-infinite bulk-bulk median interface  $z = -(d_2 + d_3)$ . Thus, the surface layer thickness is  $d_2$ ; the thickness of the bulk median layer is  $d_3$ , which appears in the

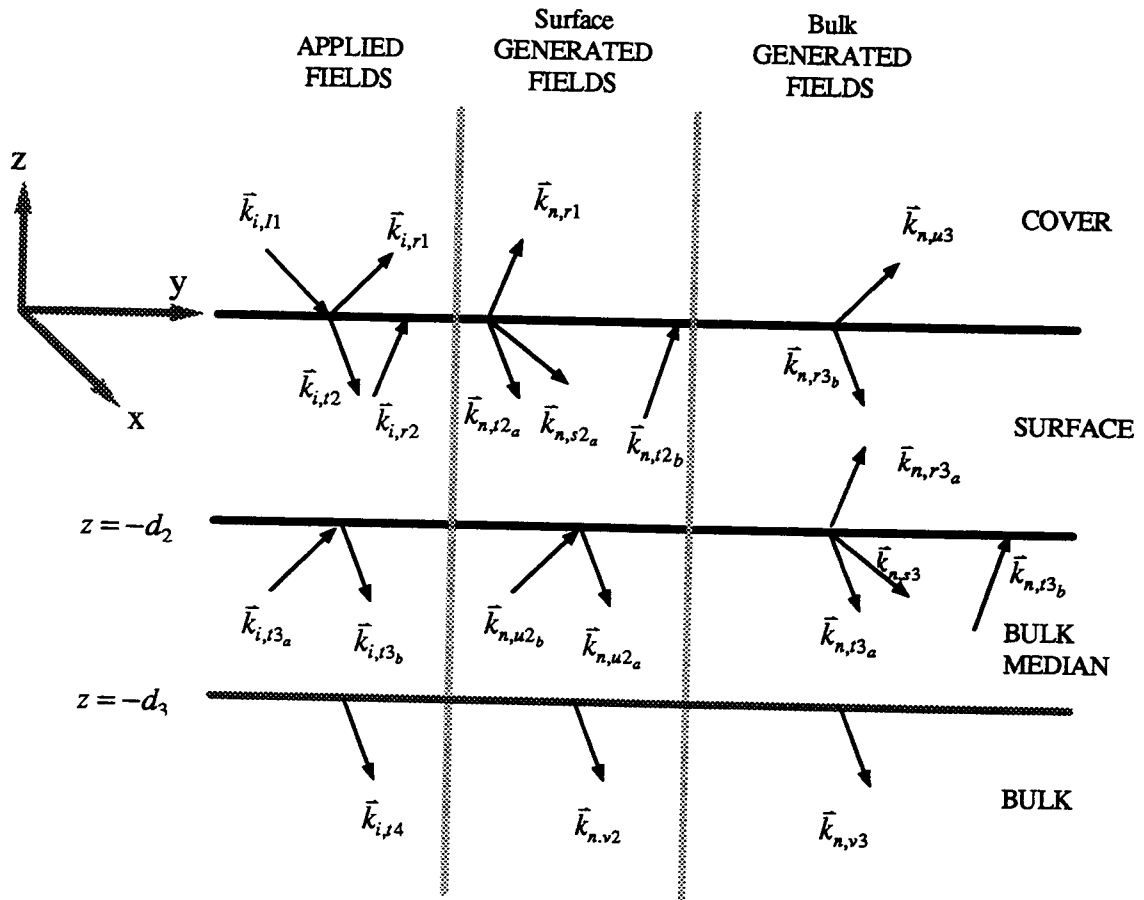


Figure 3.14: Defining the interface regions and the wavevectors of the fields.

expression for the reflected waves generated in this layer as a scaling parameter. Since the layer thickness are small compared with the wavelength of light, any linearly-related phase differences between layers are minimal. Each region is has a unique linear index of refraction indicated by  $n_1$  for the cover,  $n_2$  for the surface,  $n_3$  for the bulk median region, and  $n_4$  for the bulk region. Recall that the ratio of the (linear) indices of refraction of the bulk median and bulk,  $n_3/n_4$ , is another scaling parameter; it appears in the expression for the *amplitude* of reflected waves generated in the bulk (median).

The fields in each region propagate with associated wavevectors which are given identifying subscripts. The wavevectors and field amplitudes,  $\mathcal{E}$ , are identified with the same subscripts. These subscripts are defined by one letter describing the frequency of the wave, (a comma,) a second letter identifying the wave, followed by a number indicating the layer number in which the wave originated. (I.e.,  $\mathcal{E}_{n,r1}$  and its associated wavevector  $\vec{k}_{n,r1}$  appears at the top center portion of Figure 3.13.). The paragraphs below identify each field and their subscripts.

For the  $n-1$  incident waves, the first number is the frequency index of the wave. Since there are  $n-1$  incident waves, this first index runs from  $i = n-1$ . A comma separates the frequency index from the layer indices. The next index refers to the direction of propagation of the incident wave,  $t$  (transmitted wave) for a downwards propagating wave,  $r$  (reflected) for an upwards propagating wave. The last index is the layer index; the layer index for the incident waves represents the layer the wave is located in as shown at the far left of Figure 3.14. For instance, the applied wave '3' with frequency  $\omega_3$ , reflected from the bulk median-surface interface, has associated subscripts '3,r2'. If the process being considered was SFG, there would be two incident waves with associated subscripts '1,11' and '2,11'. The reflection of these waves in the cover layer are labeled '1,r1' and '2,r1', whereas the transmitted waves have labels for the surface layer '1,t2', '2,t2', and so forth. The wave transmitted into the bulk is given a 'v' index, for instance the wave with frequency  $\omega_3$  transmitted into the bulk has index '3,v4'. Explicitly,  $\vec{k}_{i,11}$  is the wavevector of the  $i$ -th incident wave with frequency  $\omega_i$ . It defines one of the  $n-1$  waves incident from the cover heading to the slab interface.  $\mathcal{E}_{i,11}$  is the amplitude of the field of this wave.  $\vec{k}_{i,r1}$  is the wavevector of the  $i$ -th wave reflected from the interfaces below. This wave returns to the cover (vacuum);  $\mathcal{E}_{i,r1}$  is the amplitude of the field of this wave.  $\vec{k}_{i,t2}$  is the wavevector of the  $i$ -th wave transmitted through the surface layer down towards the semi-infinite bulk;  $\mathcal{E}_{i,t2}$  is its field amplitude.  $\vec{k}_{i,r2}$  is the wavevector of the  $i$ -

$i$ -th wave which reflects from the surface-bulk interface and travels up towards the cover;  $\mathcal{E}_{i,r2}$  is the electric field amplitude of this wave.  $\vec{k}_{i,t3}$  is the wavevector of the  $i$ -th wave that is transmitted into the bulk-median.  $\mathcal{E}_{i,t3}$  is the amplitude of the field of this wave. The wave reflected from the bulk-bulk median interface has wavevector  $\vec{k}_{i,r3}$  and amplitude  $\mathcal{E}_{i,r3}$ .  $\vec{k}_{i,v4}$  is the wavevector of the  $i$ -th wave that is transmitted into the semi-infinite bulk; it has amplitude  $\mathcal{E}_{i,v4}$ . Each wavevector is described by spherical coordinates,  $\vec{k}_\lambda = (\|\vec{k}_\lambda\|, \theta_\lambda, \phi_\lambda)$ , with  $\lambda$  equal to any of the subscripts defined for the wavevectors and amplitudes above.

The generated waves have their own unique indices. As before, the generated waves, have frequency  $\omega_n$ . The layer identification number indicates *the layer number the wave is generated in*. In Bloembergen's derivation the only layer that waves are generated in was layer 2, the surface layer. Now waves are generated in the bulk median, layer 3, too. For example, the surface-generated wave parameters (which has frequency index  $n$ ) propagating into the semi-infinite bulk has subscripts ' $n,v2$ ', whereas the bulk median-generated wave propagating into the semi-infinite bulk has subscripts ' $n,v3$ '.

The waves generated in the surface region, as indicated in the center portion of Figure 3.14, are defined with the following indices to distinguish them. The wavevector of the wave generated in the surface layer, reflected into the vacuum (cover) from the interfaces below, is  $\vec{k}_{n,r1}$ . It has an amplitude  $\mathcal{E}_{n,r1}$ .  $\vec{k}_{n,t2_a}$  is the wavevector of the wave generated in surface region. It is the result of the nonlinear reflection (i.e. the 'transmitted wave') from the cover-surface interface.  $\mathcal{E}_{n,t2_a}$  is the amplitude of the field of this downwards-traveling wave. The generated wave is (linearly) reflected from the surface-bulk median interface. This reflected wave, propagating in the surface layer, has  $\vec{k}_{n,t2_b}$  as its wavevector;  $\mathcal{E}_{n,t2_b}$  is its amplitude.  $\vec{k}_{n,s2}$  is the wavevector of the 'source' wave in the surface layer. It is the effective wavevector as defined in equation (3.5). The amplitude of this wave,  $\mathcal{E}_{n,s2}$ , is

given by equations (3.12) through (3.16). When the wave generated in the surface enters the bulk median, it propagates downwards. This wave has wavevector  $\bar{k}_{n,u2_a}$  and field amplitude  $\mathcal{E}_{n,u2_a}$ . Some of this field is (linearly) reflected from the bulk-bulk median interface and propagates upwards through the bulk median. This reflected wave has  $\bar{k}_{n,u2_b}$  as its wavevector;  $\mathcal{E}_{n,u2_b}$  is its amplitude. Finally, the wave transmitted down into the semi-infinite bulk is given the wavevector  $\bar{k}_{n,v2}$  and amplitude  $\mathcal{E}_{n,v2}$ . Again, each wavevector is described by spherical coordinates:  $\bar{k}_\lambda = (\bar{k}_\lambda, \theta_\lambda, \phi_\lambda)$ , with  $\lambda$  equal to any of the subscripts defined for the wavevectors and amplitudes above.

The waves generated in the bulk median region are defined in the right-hand side of Figure 3.14. The wavevector of the wave generated in the bulk median layer is  $\bar{k}_{n,t3_a}$ . It is the result of the nonlinear reflection from the surface-bulk median interface.  $\mathcal{E}_{n,t3_a}$  is the amplitude of the field of this downwards-traveling wave. The generated wave described by  $\bar{k}_{n,t3_a}$  is (linearly) reflected from the bulk-bulk median interface. This reflected wave, propagating in the bulk median layer, has  $\bar{k}_{n,t3_b}$  as its wavevector and has amplitude  $\mathcal{E}_{n,t3_b}$ .  $\bar{k}_{n,s3}$  is the wavevector of the 'source' wave in the surface layer. It is the effective wavevector for the bulk median layer as defined in equation (3.5). The amplitude of this wave,  $\mathcal{E}_{n,s3}$ , is given by equations (3.12) through (3.16) as before with the surface layer. As the wave that is reflected off the bulk-bulk median interface travels upwards (towards the cover), it encounters the surface layer. The upwards-propagating part of this bulk median-generated wave inside the surface layer has wavevector  $\bar{k}_{n,r3_a}$  with an amplitude  $\mathcal{E}_{n,r3_a}$ . A portion of this wave is reflected off the surface-cover interface and propagates back through the surface layer, downwards. This wave has wavevector  $\bar{k}_{n,r3_b}$  and field amplitude  $\mathcal{E}_{n,r3_b}$ . The bulk median-generated wave that makes its to the cover region is referred to by its wavevector  $\bar{k}_{n,u3}$  and amplitude  $\mathcal{E}_{n,u3}$ . The wave transmitted down into the semi-infinite bulk is given the wavevector  $\bar{k}_{n,v3}$  and amplitude  $\mathcal{E}_{n,v3}$ .

Notice that only one reflection off of each interface is included. When the reflection angle is large (near grazing) this model is suspect as each layer will act as a leaky waveguide with many reflections at each interface before the light 'escapes' from the confines of the structure. This model is also limited to 'weak coupling' theory. That is the incident waves are not attenuated by the generation of a new, inhomogeneous wave.

Solutions can be found separately for the surface-generated and bulk median-generated signals as viewed in the cover region by imposing the boundary conditions on all the fields oscillating at the generated frequency (see the center and right portions of Figure 3.14). For the surface-generated wave, there are six unknown fields created from the inhomogeneous wave. Assuming a known induced nonlinear polarization (which in turn depends on the wavevectors and polarizations for all incident waves, as well as the nonlinear susceptibility of the surface medium, see (3.6)), solutions for all six fields, including the field in the cover region, can be found. Since there are two polarizations for waves in the cover region, there are two sets of solutions, one for  $S$ - and one for  $P$ -polarizations. The bulk-generated wave has similar solutions, obtained in the same manner. A complete derivation of these solutions is given in Appendix B. The approach is to treat the six homogeneous field amplitudes as a vector, and the boundary conditions at each interface as a  $6 \times 6$  matrix. The product of the homogeneous boundary condition matrix with the homogeneous field amplitude vector must yield a vector representing the satisfaction of the boundary conditions for the inhomogeneous waves. Diagonalizing the inhomogeneous boundary condition matrix yields the solutions for each of the six fields. Realizing the assumptions made, four solutions for the reflected field amplitudes arise: two for each polarization for both the surface- and bulk median-generated waves. The  $S$ -polarization reflected field amplitudes generated from the surface,  $\mathcal{E}_{nr1,S}$ , and the bulk (median),  $\mathcal{E}_{n\mu3,S}$  are (dropping the  $n$  subscript):

$$\mathcal{E}_{r1,s} = D_{r1,s}^{-1} \mathcal{E}_{s2,s} \left\{ \begin{array}{l} n_{u2} \cos \theta_{u2} \cos \phi_{u2} \begin{pmatrix} n_{s2} \cos \theta_{s2} \\ -n_{v2} \cos \theta_{v2} \end{pmatrix} (\cos \phi_{s2} - \cos \phi_{t2}) \\ + n_{u2}^2 \cos^2 \theta_{u2} \sin \phi_{u2} \left( \sin \phi_{s2} - \frac{n_{s2}}{n_{t2}} \cos \theta_{s2} \sec \theta_{t2} \sin \phi_{t2} \right) \\ + n_{v2} \cos \theta_{v2} \sin \phi_{v2} (n_{s2} \cos \theta_{s2} \sin \phi_{s2} - n_{t2} \cos \theta_{t2} \sin \phi_{t2}) \\ + i n_{u2} \cos \theta_{u2} \cos \phi_{u2} \sin \phi_{s2} \begin{pmatrix} n_{s2} \cos \theta_{s2} \\ -n_{v2} \cos \theta_{v2} \end{pmatrix} \\ - i n_{u2} \cos \theta_{u2} \cos \phi_{u2} \sin \phi_{t2} \\ \times \begin{pmatrix} n_{t2} \cos \theta_{t2} - \\ \frac{n_{s2} n_{v2}}{n_{t2}} \cos \theta_{s2} \cos \theta_{v2} \sec \theta_{t2} \end{pmatrix} \\ - i n_{u2} \cos \theta_{u2} \sin \phi_{u2} (\cos \phi_{s2} - \cos \phi_{t2}) \\ \times \begin{pmatrix} n_{v2} \cos \theta_{v2} - \frac{n_{s2} n_{v2}}{n_{t2}} \cos \theta_{s2} \cos \theta_{v2} \sec \theta_{u2} \end{pmatrix} \end{array} \right\}, \quad (3.27)$$

$$D_{r1,s} = \left\{ \begin{array}{l} n_{u2} \cos \theta_{u2} \cos \phi_{u2} (n_{r1} \cos \theta_{r1} \cos \phi_{t2} + n_{v2} \cos \theta_{v2} \cos \phi_{t2}) \\ - \begin{pmatrix} n_{t2} n_{v2} \cos \theta_{t2} \cos \theta_{v2} \\ -\frac{n_{r1} n_{u2}^2}{n_{t2}} \cos \theta_{r1} \cos^2 \theta_{u2} \sec \theta_{t2} \end{pmatrix} \\ - i n_{u2} \cos \theta_{u2} \begin{pmatrix} \cos \theta_{t2} \cos \phi_{u2} \sin \phi_{t2} \\ + n_{u2} \cos \theta_{u2} \sin \phi_{u2} \cos \phi_{t2} \end{pmatrix} \\ - i n_{r1} n_{v2} \cos \theta_{r1} \cos \theta_{v2} \begin{pmatrix} \frac{n_{u2}}{n_{t2}} \cos \theta_{u2} \sec \theta_{t2} \cos \phi_{u2} \sin \phi_{t2} \\ - \cos \phi_{t2} \sin \phi_{u2} \end{pmatrix} \end{array} \right\}, \quad (3.28)$$

$$\mathcal{E}_{u3,S} = D_{u3,S}^{-1} \mathcal{E}_{s3,S} \left\{ \begin{array}{l} \frac{n_{t3}}{n_{r3}} \cos \theta_{t3} \sec \theta_{r3} \cos \phi_{r3} \begin{pmatrix} n_{s3} \cos \theta_{s3} \\ -n_{v3} \cos \theta_{v3} \end{pmatrix} (\cos \phi_{s3} - \cos \phi_{t3}) \\ + i \frac{n_{t3}}{n_{r3}} \cos \theta_{t3} \sec \theta_{r3} \cos \phi_{r3} \begin{pmatrix} n_{s3} \cos \theta_{s3} \sin \phi_{s3} \\ -n_{v3} \cos \theta_{v3} \sin \phi_{s3} \\ -n_{t3} \cos \theta_{t3} \sin \phi_{t3} \\ + \frac{n_{s3} n_{v3}}{n_{t3}} \cos \theta_{s3} \cos \theta_{v3} \\ \times \sec \theta_{t3} \sin \phi_{t3} \end{pmatrix} \end{array} \right\}, \quad (3.29)$$

$$D_{u3,S} = \left\{ \begin{array}{l} n_{t3} \cos \theta_{t3} \begin{pmatrix} n_{s3} \cos \theta_{s3} \\ -n_{v3} \cos \theta_{v3} \end{pmatrix} \begin{pmatrix} \cos \phi_{t3} \sin \phi_{r3} \\ -\frac{n_{t3}}{n_{r3}} \cos \theta_{t3} \sin \phi_{t3} \sec \theta_{r3} \cos \phi_{r3} \end{pmatrix} \\ -i \frac{n_{u3} n_{v3}}{n_{r3}} \cos \theta_{u3} \cos \theta_{v3} \sec \theta_{r3} \sin \phi_{t3} \end{array} \right\}. \quad (3.30)$$

The  $P$ -polarization reflected field amplitudes generated from the surface,  $\mathcal{E}_{nr1,P}$ , and the bulk (median),  $\mathcal{E}_{nu3,P}$  are (dropping the  $n$  subscript):

$$\mathcal{E}_{r1,P} = D_{r1,P}^{-1} \mathcal{E}_{s2,P} \left\{ \begin{array}{l} n_{u2} \cos \theta_{u2} \cos \phi_{u2} \left[ (\mathcal{E}_{B2} - \mathcal{E}_{D2}) \sin \phi_{s2} - \left( \mathcal{E}_{B2} \frac{\cos \theta_{t2}}{n_{t2}} - \mathcal{E}_{D2} \sin \phi_{s2} \right) \sin \phi_{t2} \right] \\ - \cos^2 \theta_{u2} \sin \phi_{u2} [\mathcal{E}_{B2} (\sin \phi_{s2} - \cos \phi_{t2})] \\ + n_{u2}^2 \sin \phi_{u2} [\mathcal{E}_{D2} (\cos \phi_{s2} - \cos \phi_{t2})] \\ - i n_{u2} \cos \theta_{u2} \cos \phi_{u2} [(\mathcal{E}_{B2} - \mathcal{E}_{D2}) (\cos \phi_{s2} - \cos \phi_{t2})] \\ - i \cos^2 \theta_{u2} \sin \phi_{u2} [\mathcal{E}_{B2} \sin \phi_{s2} - \mathcal{E}_{D2} \sec \theta_{t2} \sin \phi_{t2}] \\ + i n_{u2}^2 \sin \phi_{u2} \left[ \mathcal{E}_{D2} \sin \phi_{s2} - \mathcal{E}_{B2} \frac{\cos \theta_{t2}}{n_{t2}} \sin \phi_{t2} \right] \end{array} \right\}, \quad (3.31)$$



$$D_{r1,P} = \left\{ \begin{aligned} & n_{u2} \cos \theta_{r1} (n_{t2} \cos \theta_{u2} \cos \phi_{u2} \sec \theta_{t2} \sin \phi_{t2} + n_{u2} \sin \phi_{u2} \cos \phi_{t2}) \\ & + \frac{n_{r1} n_{u2}}{n_{t2}} \cos \theta_{t2} \cos \theta_{u2} \cos \phi_{u2} \sin \phi_{t2} \\ & + n_{r1} \cos^2 \theta_{u2} \sin \phi_{u2} \cos \phi_{t2} \\ & + i n_{u2} \cos \theta_{u2} (n_{r1} + \cos \theta_{r1}) \cos \phi_{u2} \cos \phi_{t2} \\ & - i \frac{n_{r1} n_{u2}^2}{n_{t2}} \cos \theta_{t2} \sin \phi_{u2} \sin \phi_{t2} \\ & - i n_{t2} \cos \theta_{r1} \cos^2 \theta_{u2} \sec \theta_{t2} \sin \phi_{u2} \sin \phi_{t2} \end{aligned} \right\}, \quad (3.32)$$

$$\mathcal{E}_{u3,P} = D_{u3,P}^{-1} \mathcal{E}_{s3,P} \left\{ \begin{aligned} & \frac{\mathcal{E}_{B3}}{4} [\cos(2\theta_{t3} - \phi_{t3}) + \cos(2\theta_{t3} + \phi_{t3}) - 2\cos \phi_{t3}] \\ & + \frac{\mathcal{E}_{B3}}{2} [\cos(\theta_{t3} - \phi_{t3}) + \cos(\theta_{t3} + \phi_{t3})] \\ & + \frac{n_{t3}}{2} (\mathcal{E}_{B3} - \mathcal{E}_{D3}) [\cos(\theta_{t3} - \phi_{t3}) - \cos(\theta_{t3} + \phi_{t3})] \\ & - \frac{i}{2} (\mathcal{E}_{B3} - \mathcal{E}_{D3}) [\sin(\theta_{t3} - \phi_{t3}) - \sin(\theta_{t3} + \phi_{t3})] \\ & - \frac{i}{2} \mathcal{E}_{D3} [\sin(\theta_{t3} - \phi_{t3}) - \sin(\theta_{t3} + \phi_{t3}) + 2n_{t3}^2 \sin \phi_{t3}] \end{aligned} \right\}, \quad (3.33)$$

$$D_{u3,P} = \left\{ \begin{aligned} & \left( n_{u3} \cos \theta_{t3} \cos \phi_{r3} \cos \phi_{t3} \right. \\ & \left. - \frac{n_{t3} n_{u3}}{n_{r3}} \cos \theta_{r3} \sin \phi_{r3} \sin \phi_{t3} \right) (\cos \theta_{t3} - n_{t3}) \\ & - i \left( n_{t3} \cos \theta_{u3} \cos \phi_{r3} \sin \phi_{t3} \right. \\ & \left. - n_{r3} \cos \theta_{u3} \cos \theta_{t3} \cos \phi_{t3} \sec \theta_{r3} \sin \phi_{r3} \right) (\cos \theta_{t3} - n_{t3}) \end{aligned} \right\}. \quad (3.34)$$

In the expressions for the reflected fields, the following definitions were made:

$$\phi_{u2} \equiv \frac{\omega_n}{c} n_{u2} d_3 \cos \theta_{u2}, \quad (3.35)$$

$$\phi_{s2} \equiv \frac{\omega_n}{c} n_{s2} d_2 \cos \theta_{s2}, \quad (3.36)$$

$$\phi_{v2} \equiv \frac{\omega_n}{c} n_{v2} (d_2 + d_3) \cos \theta_{v2}, \quad (3.37)$$

$$\phi_{t2} \equiv \frac{\omega_n}{c} n_{t2} d_2 \cos \theta_{t2}, \quad (3.38)$$

$$\Phi_2 \equiv - \frac{\left( \bar{k}_{n,t2} - \left( \sum_{i=1}^{n-1} \bar{k}_{i,t2} \right) \right) \cdot \bar{k}_{n,t2}}{|\bar{k}_{n,t2}|} \cdot \frac{d_2}{2 \cos \theta_{n,t2}} \equiv - \frac{(\bar{k}_{n,t2} - \bar{k}_{n,s2}) \cdot \bar{k}_{n,t2}}{|\bar{k}_{n,t2}|} \cdot \frac{d_2}{2 \cos \theta_{n,t2}}, \quad (3.39)$$

$$\phi_{u3} \equiv \frac{\omega_n}{c} n_{u3} d_2 \cos \theta_{u3}, \quad (3.40)$$

$$\phi_{s3} \equiv \frac{\omega_n}{c} n_{s3} d_3 \cos \theta_{s3}, \quad (3.41)$$

$$\phi_{v3} \equiv \frac{\omega_n}{c} n_{v3} (d_2 + d_3) \cos \theta_{v3}, \quad (3.42)$$

$$\sin \phi_{t3} \equiv \frac{\omega_n}{c} n_{t3} d_3 \cos \theta_{t3}, \quad (3.43)$$

$$\Phi_3 \equiv - \frac{\left( \bar{k}_{n,t3} - \left( \sum_{i=1}^{n-1} \bar{k}_{i,t2} \right) \right) \cdot \bar{k}_{n,t3}}{|\bar{k}_{n,t3}|} \cdot \frac{d_3}{2 \cos \theta_{n,t3}} \equiv - \frac{(\bar{k}_{n,t3} - \bar{k}_{n,s3}) \cdot \bar{k}_{n,t3}}{|\bar{k}_{n,t3}|} \cdot \frac{d_3}{2 \cos \theta_{n,t3}}. \quad (3.44)$$

For *P*-polarizations, the additional definitions are

$$\mathcal{E}_D \equiv (\cos\theta_k, \sin\alpha) \quad (3.45)$$

and

$$\mathcal{E}_B \equiv \left( \sin\theta_k, \cos\alpha \frac{(n_i^2 - n_s^2)}{n_i^2} \right). \quad (3.46)$$

The solutions presented here are quite general: they were found without exact knowledge of the induced nonlinear polarization. Hence, the solutions obtained are independent of the OM process (the order of the nonlinear polarization) and are general with respect to the incident wave parameters. Bloembergen, in his model, approximated the thickness of the surface region as small compared with the generated wavelength. This approximation is not done here. Moreover, the exact phase-matching expression is retained. The phase-mismatch information for the surface-generated wave is contained in the argument  $\Phi_2$  of the *sinc*  $\Phi_2$ ;  $\Phi_2$  is referred to as the *surface wave phase-mismatch parameter*. In addition, combinations of  $\phi_{i2} - \phi_{s2}$  and  $\theta_{i2} - \theta_{s2}$  can be thought of as constituents of the *surface wave phase-mismatch factor* to the generated field amplitude. The *surface wave phase-mismatch parameter* represents difference between the generated wavevector,  $\bar{k}_{n,i2_a}$ , and the effective 'source' wavevector,  $\bar{k}_{n,s2}$ , as defined in (3.5). Likewise, there is a *bulk wave phase-mismatch parameter*,  $\Phi_3$ , and a *bulk wave phase factor*. This phase-mismatch information is paramount when comparing the phase-matched surface signal with the (possibly) phase-mismatched bulk signal. The

approximations made in the Bloembergen model concerning phase-matched conditions (small phase-mismatch) are not appropriate for this model in which phase-mismatch may be large.

A very useful form of the reflected surface and bulk field amplitudes (see Appendix B) may be written as

$$\mathcal{E}_{r1,\sigma} = \mathcal{E}_{BC,r1,\sigma} \mathcal{E}_{s2,\sigma} = \mathcal{E}_{BC,r1,\sigma} \left( \frac{-4\pi P_\sigma}{n_{t2}^2 - n_{s2}^2} \right) \sec\theta_{t2} \operatorname{sinc}\Phi_2 e^{-i\Phi_2} \quad (3.47)$$

and

$$\mathcal{E}_{u3,\sigma} = \mathcal{E}_{BC,u3,\sigma} \mathcal{E}_{s3,\sigma} = \mathcal{E}_{BC,u3,\sigma} \left( \frac{-4\pi P_\sigma}{n_{t3}^2 - n_{s3}^2} \right) \sec\theta_{t3} \operatorname{sinc}\Phi_3 e^{-i\Phi_3}. \quad (3.48)$$

This is done by defining  $\mathcal{E}_{BC,r1,\sigma}$  and  $\mathcal{E}_{BC,u3,\sigma}$  as the factors found by matching the boundary conditions of the layered structure in Figure 3.14. That is,  $\mathcal{E}_{BC,r1,\sigma}$  and  $\mathcal{E}_{BC,u3,\sigma}$  are the contents (including the denominators) of the curly brackets in (3.27) through (3.34). Here  $\sigma$  indicates either the *S* or *P* polarization. Clearly, the surface and bulk reflected fields have a sinusoidal and a *sinc* dependence on the phase-mismatch parameter. In this form, the  $\sec\theta$  dependence is readily apparent. Also, the  $(n_t^2 - n_s^2)^{-1}$  factor, which constitutes an additional dependence on phase-mismatch, is clearly separated from the boundary layer aspects of the reflected amplitudes. Using this form of the reflected field amplitudes, the various dependencies can be investigated individually. However, for perfect phase-matching the product  $\mathcal{E}_{BC}(n_t^2 - n_s^2)^{-1}$  as  $n_s \rightarrow n_t$  yields a finite result. Hence, for phase-matched conditions this product must be treated as a whole. Additionally, since  $P_S$  and  $P_P$  are taken to be equal, all the polarization dependence is contained in the  $\mathcal{E}_{BC,r1,\sigma}$  and  $\mathcal{E}_{BC,u3,\sigma}$  factors via  $\mathcal{E}_D$  and  $\mathcal{E}_B$  (see (3.45) and (3.46)).

The next step in the prediction of RISOM experiments is to assume a nonlinear susceptibility and polarizations and wavevectors for the incident waves, calculate the induced nonlinear polarization, from this explicitly calculate the field amplitudes from (3.27) through (3.44), and compare the ratio to surface-generated to bulk median-generated amplitudes (in the cover region). Repeated calculations are made so as to map out the parameter space looking for favourable conditions for RISOM.

### 3.4.2 Summary and Discussion of the Macroscopic RISOM Model

There are two clear predictions made by the macroscopic model. First is the enhancement of the surface over the bulk signal due to phase-mismatch. If dispersion of the surface and bulk (median) are different, the phase-mismatching parameter will be different in the two layers. For a generated wave, the angle at which its phase-mismatch is minimum will be the angle it will 'reflect' into the cover region at. Due to a difference in dispersion of the surface and bulk median layers, the 'reflected' surface and bulk waves will not be collinear. If the bulk median and surface have the same non-zero dispersion, it is possible that even though the surface and bulk waves 'reflect' at the same angle, the phase-mismatch of the surface is small, but the phase-mismatch for the bulk median is large. Hence, the surface signal would be large compared to the bulk signal. The second prediction involves the separation of the surface and bulk signals due to dissimilar NLBAs. For *P*-waves, the NLBA for perfect phase-matching ( $n_2 = n_{s2}$ ) occurs at  $\theta_2 = (\pi - \alpha_2)/2$ . Similarly, for the bulk region  $\theta_2 = (\pi - \alpha_3)/2$ . Any difference in dispersion leads to the condition where the NLBA for the bulk is unequal to the NLBA for the surface. An experiment designed for *P*-waves reflecting from the structure at the NLBA of the bulk will yield a small, but non-zero, signal from the surface. This is possible since the NLBA for the surface is not that the same angle as the experimental angle (the NLBA of the bulk). The degree of enhancement, the sensitivity of angle, degree of dispersion, etc., for these two predictions is, however, not easily obtained without extensive numerical analysis. Furthermore, it is not trivial to locate other possible enhancements that are due

to effects of the linear and nonlinear Fresnel coefficients. This is more efficiently done by numerical calculations using the RISOM field expressions. The object of the remainder of this chapter is to quantify the enhancements.

This model predicts a  $\sec \theta_{r,l}$  behaviour on the amplitude of the generated waves. Thus it is expected that the surface reflected wave should increase in intensity as the angle increases to near grazing angles. The bulk-generated wave amplitudes are shown to have a more complicated angular dependence due to the added Fresnel coefficients introduced by the surface-bulk median interface. This angular dependence is less strong than  $\sec \theta_{r,l}$ , as is seen in the numerical calculations to follow. From this observation, it is expected that the surface reflected wave should increase in intensity as the angle increases to near grazing angles more rapidly than the bulk-generated signal.

### 3.4.3 Prelude to Numerical Analysis of the RISOM expressions

The remainder of this chapter is devoted to mapping out the parameter space of the model's predictions, specifically where ISOM is likely. The analysis is divided into two parts: analysis of collinear phase-matched geometries and analysis of non-collinear phase-matched geometries. Collinear phase-matching is appropriate for OM in dispersionless materials and for harmonic generation. Non-collinear phase-matching is where the linear dispersion of materials is important or where the incident wavevectors are not completely degenerate (such as in SFG with waves of differing frequencies or non-degenerate SHG with fundamental waves incident at different angles). Here the angular dependencies for the two polarization states for SHG (where dispersion is very important, but phase-matching is automatic) and CARS-like FWM (where dispersion is less important, but phase-matching is critical) are specifically considered. Both polarization states,  $S$  and  $P$ , for both types of geometries are explored. In general, it is found that improvement in the surface-to-bulk signal ratio is possible for  $S$ -polarization when the reflection angle is large; for  $P$ -waves, enhancement is done by making use of the nonlinear Brewster's angle. It is

found that the results for enhancing the signal-to-noise ratio for collinear phase-matching are the same for the non-collinear case. However, enhancement is generally better if the surface and bulk dispersions are dissimilar. In addition, it is shown that there is an angular separation of bulk and surface signals due to dispersion (for non-collinear phase-matching). When the index of refraction is different in the surface and bulk for a given frequency, the optical process in the surface may be phase-matched, whereas the bulk may not be. Hence, waves generated in the surface and bulk travel in (slightly) different directions. With this prelude to the numerical analysis, the development continues.

### **3.5 Introduction to the Numerical Analysis of the RISOM Expressions**

The analysis of the RISOM expressions are particularly involved. Before any calculations are presented, explanatory information is introduced concerning the calculations and interpretive tools used in the analysis. It is also useful to be prepared as to the direction that the analysis takes. As the analysis proceeds, more complexity is added to the picture. These complexities build on each other making interpretations increasingly more difficult. The logic in the analysis sections is to begin with the most fundamental 'experiments' and to separately add each complexity to the overall picture. This introduction section is devoted to outlining the analysis made of the RISOM expressions.

The contributions of the surface and bulk regions to the electric fields were numerically calculated for a variety of cases. These calculations are based on the layered structure shown in Figure 3.13 and the RISOM expressions (3.27) through (3.46). Particular attention was paid to conditions which might enhance the surface over the bulk contribution. The analysis is divided into two parts: collinear geometry and non-collinear geometry. In each case general and specific OM processes are considered for both

polarizations and in media with and without dispersion. When appropriate, RISOM experiments are suggested and discussed. The next paragraphs will briefly outline the directions the analysis takes.

Calculations with collinear optical mixing are used to show the most fundamental angular dependence for the bulk and surface contributions. These results hold for *all* collinear mixing processes including degenerate  $n$ -th harmonic generation ( $n$ HG), and collinear general TWM and FWM. Any optical mixing done in dispersionless media is automatically phase-matched in collinear geometry. The first analysis of general collinear optical mixing is done in dispersionless media for  $S$ -polarizations. A RISOM experiment which shows reasonably good enrichment of the surface contribution to the total signal is predicted for near-grazing angles of incidence. Also, it is shown that the phase of the surface and bulk waves are always virtually constant. This is due to the coherence of the nonlinear process.

The next part of the analysis considers  $S$ -wave  $n$ HG processes in media with dispersion. General OM processes are not covered in this analysis since the mismatch is extremely large; experiments involving general OM processes are best considered in non-collinear geometries. The inclusion of dispersion in collinear  $n$ HG processes will not change the angular behaviour bulk and surface contributions, however it is shown that the overall intensity drops because of a phase-mismatch between the induced polarization and the generated wave. This is due to the dispersion between the fundamental and harmonic waves. When the dispersions of the surface and bulk are different, the wave generated in the medium with the smallest dispersion contributes more to the overall observed intensity. When the dispersion is greater, so is the phase-mismatch, and the intensity generated in the medium is smaller.

Analysis of  $P$ -waves is more complicated than that of  $S$ -polarization.  $P$ -waves may or may not exhibit a nonlinear Brewster's angle (NLBA). The analysis of  $P$ -waves separates



these two possibilities. First *P*-polarization collinear *n*HG in dispersionless media when there is no NLBA is scrutinized. It is discovered that much of the angular dependence of the generated field amplitudes for *P*-polarization collinear *n*HG is the same as for *S*-polarizations. A fair surface-to-bulk signal ratio can be obtained when performing experiments at near-grazing angles of incidence. Much of the analysis (for *P*-waves) carries over (from the *S*-wave analysis), however, the phase of the surface and bulk waves can be different depending on the direction of the induced polarization and mixed wave.

Analysis continues with *P*-wave collinear *n*HG processes with a NLBA. In order to analyze the NLBA, the generated frequency must be specified. Thus the analysis is done for SHG. The frequency generated is chosen to be 532 nm. Initially the polarization directions of the surface region and the bulk are considered equal. Again, both surface and bulk media are considered dispersionless. The main conclusion is that, if the media are dispersionless, there is no means to separate the bulk and surface signals in this geometry.

Further analysis of the NLBA for *P*-waves in a collinear SHG process continues by including dispersion in the surface and bulk median. If the dispersions of the bulk and surface are different, the NLBAs will differ. Even with small dispersions, the disparity between NLBAs can be experimentally observable. Unique bulk and surface NLBAs offer another way to emphasize one signal over the other. For instance, it is shown that when the bulk experiences a NLBA this signal is zero, but since the NLBA of the surface is different, there is still a surface signal. This experiment offers excellent surface-to-bulk signal ratios, but at the expense of the detectable signal intensity.

As foreshadowed in the analyses of collinear reflection geometry OM, non-collinear geometries hold the key to successful RISOM experiments. From the series of collinear calculations done on media with dispersion, it was seen that dispersion adds mismatch. Different mismatches of the OM process in surface and bulk median media produce ways

to separate the contributions from these layers. Another way to introduce phase-mismatch is to slightly change the direction of one (or more) of the incident wavevectors from the phase-matched case. This introduces a mismatch that is to the discretion of the experimenter. There are two experimental designs for non-collinear geometries. In either case, the incident frequency source is split into two waves which are incident on the sample with slightly different angles of incidence. These two experimental situations can be superimposed comprising a more general 'experiment'. The phase-mismatch can be introduced either by selecting different polar angles, leaving the azimuthal angles similar, or by the converse. The other reason for considering non-collinear geometries is that for processes that are not degenerate in frequency done in media with dispersion, non-collinear geometries are the only way to minimally phase-mismatch (and maximize the mixed intensity). There are two main conclusions that will come of the analyses of non-collinear geometries. First, when the media have no dispersion, the surface and bulk waves are 'emitted' from the structure at the same angle. They can be separated in intensity by phase-mismatching the mixing process in one layer more than the other. It is found that there are restrictions on the parameter space which allows the experimenter to separate these signals. The other conclusion is that with the inclusion of dispersion, the bulk and surface waves are separated in angle.

Actual experiments do not use plane waves as the RISOM model assumes. The incident light would be in the form of Gaussian beams. Furthermore, these beams would be focused onto the sample as to maximize the induced polarization, and hence, mixed signal. When appropriate, analyses of Gaussian beam experiments are performed. These approximate Gaussian beam analyses give results representative of real laboratory experiments. As in the collinear case, each possibility of mixing type (SHG, TWM, FWM) is investigated separately. Each process is divided into analysis of polarization, dispersion, and experimental design as described below.

First in the series of non-collinear reflection OM analysis is an experiment that introduces a mismatch via breaking the degeneracy in polar angles. To simplify analysis, only SHG is considered; this provides an investigation of the most fundamental aspects of RISOM in phase-mismatched non-collinear geometry. Analysis for *S*-wave and *P*-wave (without a NLBA) non-collinear SHG is done simultaneously since the model predicts both polarizations are similar for collinear geometry experiments. Numerical calculations are used to investigate effects of two variables. These two variables are the (average) polar angle of the incident waves and the difference in the polar angles of the waves (with all azimuthal angles equal). Because dispersion is ignored, the bulk and surface waves 'reflect' off the surface at the same angle. The angular dependencies for the surface and bulk are found to be different and complicated. This affects the surface-to-bulk ratio in a non-trivial way. There is a way to separate the bulk and surface signals via phase-mismatch. Certain sets of conditions give the surface signal a smaller sensitivity to mismatch than bulk signal. As the mismatch is increased, the bulk intensity drops rapidly leaving predominately surface contributions to the total signal. Estimates for a Gaussian beam SHG experiment are made; they show that for certain parameters, the SHG beam has a profile that is, in some places, enriched in surface contribution. The parameter space yielding these RISOM conditions is scrutinized in detail. It is found that for any set of indices and layer thicknesses, it is likely that there exists a set of incidence angles for which intensity of the bulk wave is very sensitive to phase-mismatch. The surface signal is only moderately sensitive to mismatching.

Using the azimuthal angles to achieve phase-mismatched non-collinear SHG was the other possible RISOM experiment that can be performed. Dispersion in the media is, initially, not included. Analysis is similar to the polar angle experiment, however there are different conclusions. It is found that the bulk wave is rather insensitive to the azimuthally associated phase-mismatch as compared to the dependence of the surface wave. Thus, for small phase-mismatch the total signal mismatch dependence is dominated by the surface dependence. Unfortunately, this leaves no way to attenuate the bulk signal over the

surface wave. The converse is possible, however, providing a way to observe the isolated bulk signal. From observations like this, bulk nonlinear susceptibility tensor elements may be examined.

The analysis of the polar and azimuthal difference angle experiments are done for *S*- and *P*-polarization (when no NLBA is present) in parallel. In the absence of dispersion the NLBAs for the surface and bulk are equal and no segregation of the surface and bulk contributions can be made. It will be discovered that separation is only possible when there exists a phase-mismatch. The first experiment considered, in which a phase-mismatch was introduced by breaking the polar angle degeneracy of the incident waves, may be used to improve the surface-to-bulk signal ratio. If the NLBA happens to be near the angle at which the bulk is most sensitive to mismatches, then the mismatch can be used to eliminate the bulk signal over the surface contribution. This, unfortunately, comes at the expense of the total signal strength.

After this lengthy examination of non-collinear SHG in dispersionless media, a general discussion ensues concerning the influences of dispersion in these processes. The analysis includes both polarization cases and lays the foundation for further discussion of TWM and FWM processes. Essentially, the effects of dispersion in the collinear case is superimposed on the intentionally introduced phase-mismatch of the non-collinear geometry. It is shown that the inclusion of dispersion does not affect the angular dependence of the harmonic waves; this includes the NLBAs. Also the waves 'reflect' from the structure at the same angles. By the arguments presented in earlier analyses, it is possible, using non-collinear geometry, to almost completely separate the surface and bulk signals in intensity. When the indices of the surface and bulk are different via effects of disparate dispersions, the phase-mismatch of the surface and bulk differ. If the surface mismatch can be chosen to be small and the bulk large, the surface-to-bulk signal ratio may be quite favourable. This is not a very exciting analysis, but it lays the groundwork for discussions of higher order mixing.

General non-collinear TWM in dispersionless media is the subject of the next section of the analysis. It is argued that TWM is similar to SHG as previously examined. The difference is that the process is minimally phase-mismatched (or has zero phase-mismatch) when the incident waves are *not* collinear. An expression is quoted for the matching angles. As with SHG, altering the directions of the incident waves from their minimally phase-mismatched positions introduces an experimentally controlled phase-mismatch that can be used to separate the surface and bulk contributions. This is possible when the bulk wave is more sensitive to mismatch than the surface.

Fully breaking the degeneracy (direction and frequency) of TWM processes in the presence of dispersion actually leads to a finite angular separation between the surface and bulk waves. This argument is substantiated by considering that the generated waves ‘reflect’ at the angle for which the phase-mismatch is minimum (given the incident wavevectors and differences in dispersion). Expressions for the bulk and surface ‘reflection’ angles are given. Specific examples of plane wave SFG processes and SFG using Gaussian beams are investigated. The mixed Gaussian beam ‘experiment’ shows an angular profile which has a bright spot (the surface signal) surrounded by a much larger dim halo (the bulk signal). Use of phase-mismatching can improve the signal-to-bulk ratio as discussed. If a NLBA exists, then at least for *P*-polarization, further enhancements in the signal ratio can be made.

As with TWM, an analysis for general non-collinear FWM in media without dispersion is performed. It is found that the addition of a third incident wave makes it easier to introduce a phase-mismatch in the bulk and not the surface. Finally, nonlinear geometry FWM in media with dispersion is discussed. Results are seen to be similar to the TWM case with dispersion, however angular separation between the surface and bulk is more

pronounced. Careful analyses using Gaussian beams are given; they show that the mixed beam has an intensity profile. The surface and bulk contributions can be viewed more or less individually.

### 3.5.1 Definitions and Common Terms Used in the Analysis

Throughout this analysis there are certain terms and definitions that should be introduced or reviewed. All interpretations originate from the field expressions, or ‘RISOM expressions’, equations (3.27) through (3.46). These field expressions represent the bulk median-generated and surface-generated fields for *S*- and *P*-polarizations. The field expressions are also referred to as the *nonlinear reflection coefficients* (NLRC). The NLRCs are multiplied by constant amplitude, polarization, and phase factors to form the *complete* electric field vector. (See (3.12) and (3.13).) The constant amplitude and phases are neglected in the analysis; the NLRCs alone carry all the dependence needed to predict RISOM conditions. Notice the NLRCs carry both a real part and an imaginary part. Thus they contribute an angular and an index of refraction dependence on the *complete* amplitude *and* phase of the generated waves. The *complete* surface and bulk median field amplitudes are written as  $|E_{Surface}|$  and  $|E_{Bulk}|$ , respectively. Note that the signal-to-bulk signal ratio is

$$\frac{|E_{Surface}|}{|E_{Bulk}|}.$$

Note that this ratio becomes infinite as the bulk signal approaches zero. Thus, in numerical calculations, this ratio is often avoided. The intensities of these waves are, of course, the complex squares of the field expressions,  $|E_{Surface}|^2$  and  $|E_{Bulk}|^2$ , and the *total* signal intensity is  $|E_{Surface} + E_{Bulk}|^2$ . The phase of the waves generated in the surface and bulk median layers are defined in the usual way:

$$\tan \phi_{E_{Surface}} = \text{Im}[E_{Surface}] / \text{Re}[E_{Surface}]$$

and

$$\tan \phi_{E_{Bulk}} = \text{Im}[E_{Bulk}] / \text{Re}[E_{Bulk}].$$

The direction of each wave, incident or generated, is characterized by its wavevector. As earlier stated each incident (applied) wave has wavevector described by its spherical coordinates  $\bar{k}_{i,I} = [\bar{k}_{i,I}, \theta_{i,I}, \phi_{i,I}]$ . For collinear geometries, all wavevectors have the same polar and azimuthal angles.

### 3.5.2 Parameter Settings

Throughout most of the analysis of the RISOM expressions, calculations are done with a consistent set of fixed parameters. These parameters include the indices of refraction of the layers,  $n_1$ ,  $n_2$ ,  $n_3$ , and  $n_4$ ; the thicknesses of the surface layer,  $d_2$ ; and the thickness of the bulk median layer,  $d_3$ . The index of refraction of the cover region is the index of the vacuum,  $n_1 = 1.00$ . The index of the surface, bulk, and bulk median are seen as scaling parameters. They are set to: surface index  $n_2 = 1.50$ , bulk median index  $n_3 = 1.55$ , and bulk index  $n_4 = 1.56$ . These media are considered dispersionless unless otherwise stated. The thicknesses  $d_2$  and  $d_3$ , the layer thicknesses, are considered small compared to the wavelength of the light (typically 532 nm). The layer thicknesses are both set at 5 nm. In each of these calculations the nonlinear susceptibility is considered a scalar and constant. There are no preferred directions or frequency dependencies for the induced polarization. This way only the influences of the NLRCs are investigated. The symmetry of the media and their frequency dependencies are factored out of the NLRCs. Normally *complete* fields are products of the frequency and symmetry independent NLRCs and the nonlinear polarizabilities which carry these dependencies. These dependencies can be included separately from the NLRCs if required.

### 3.5.3 How to Interpret Analysis Plots

In Chapter 2, there were two ratios that were introduced as interpretive tools in the evaluation of ISOM situations. These ratios were called the direct ratio (DR) which appears as

$$DR \equiv \frac{|E_{Surface}|^2}{|E_{Surface}|^2 + |E_{Bulk}|^2}$$

and the cross ratio (CR) which was defined by

$$CR \equiv \frac{E_{Bulk}E_{Surface}^* + c.c.}{|E_{Bulk}|^2}.$$

The DR and CR are the same ratios used in the analysis in Chapter 2. The reader is encouraged to review Section 2.5.2 which discusses how the DR and CR are to be interpreted. While reviewing, the reader will recall Figure 2.11. It represents a plot of the total combined surface and bulk intensity versus the surface field. The figure shows how the total intensity changes for a fixed bulk field, but increasing surface field amplitude. Figure 2.12 and Figure 2.13 show the DR and CR for the case of Figure 2.11. Although in the analysis that follows both the surface and bulk fields vary as the experimental parameters are changed, the DR and CR are interpreted in the same manner.

The DR is a better interpretive tool than the surface-to-bulk amplitude ratio since the DR never diverges. The direct ratio is an appropriate tool to use when the phase of the surface and bulk (complete) fields are approximately equal. Although the CR is not



bounded, it can be used to investigate the relative phases of the surface and bulk waves. When, for example, the surface and bulk fields are of the same magnitude the CR is +2 if the phases are of the same sign, -2 if they are opposite in sign. Generally, when the phases of waves are of opposite sign the CR is negative.

### 3.6 Analysis of RISOM in Collinear Geometry

Collinear geometry is appropriate for harmonic generation and general optical mixing in dispersionless media. Since expressions (3.27) through (3.46) make no assumptions about the induced polarization, it is not necessary to know the precise OM process explicitly. Hence, second harmonic generation (SHG), third harmonic generation (THG), (for any integer,  $n$ )  $n$ -th harmonic generation ( $n$ HG), and any mixing process in a dispersionless material can be investigated. In Section 3.6.1 calculations with phase-matched collinear mixing in the model structure with no dispersion is used to find the angular dependence of the nonlinear reflection coefficients for the bulk and surface contributions. Determination of  $S$ -wave NLRCs for collinear optical mixing in dispersionless materials is done first in Section 3.6.1.1. In Section 3.6.1.2, dispersion is added to the  $S$ -wave  $n$ HG picture. The analysis of  $P$ -waves is divided into two cases. First  $P$ -polarization collinear  $n$ HG in dispersionless media when there is no NLBA is scrutinized. Analysis continues with  $P$ -wave collinear  $n$ HG processes with a NLBA. In order to analyze the NLBA, the generated frequency must be specified. Thus this analysis is done for SHG generating 532 nm. The  $P$ -wave analyses are done in Section 3.6.1.3 for media without dispersion, and with dispersion in Section 3.6.2.

#### 3.6.1 Collinear General Optical Mixing

In studying collinear geometry first, the most fundamental angular dependence of the nonlinear reflection coefficients is examined. This dependence includes the  $\sec \theta$  and the parts of  $\mathcal{E}_{BC}$  that remain when  $n_t = n_s$ . From (3.47) and (3.48) it is clear that these

dependencies are always present whether the process is phase-matched or not. Hence, these dependencies will always be present even in the complicated experiments introduced later.

### 3.6.1.1 S-polarization NLRC for Collinear Optical Mixing in Dispersionless Media

Figure 3.15 shows the angular dependence of the reflected field amplitude for the generated surface wave. Since the reflected surface and bulk waves propagate collinearly, all subscripts are dropped; all waves propagate at the same polar angle,  $\theta$ , measured with respect to the normal. The azimuthal angles,  $\phi$ , are set to zero. Figure 3.16 shows the reflected field amplitude for the generated bulk field, whereas Figure 3.17 shows the DR versus incident angle,  $\theta$ . Due to the  $\sec \theta$  behaviour, the fields are stronger at high angles. Moreover, the DR plot shows the enrichment of the surface contribution to the total signal at near-grazing angles. This is due to the complicated influences of  $\mathcal{E}_{BC,2}$  and  $\mathcal{E}_{BC,3}$  on  $\theta$ . Since the phases of the surface and bulk are uninteresting, the CR for this experiment is not shown.

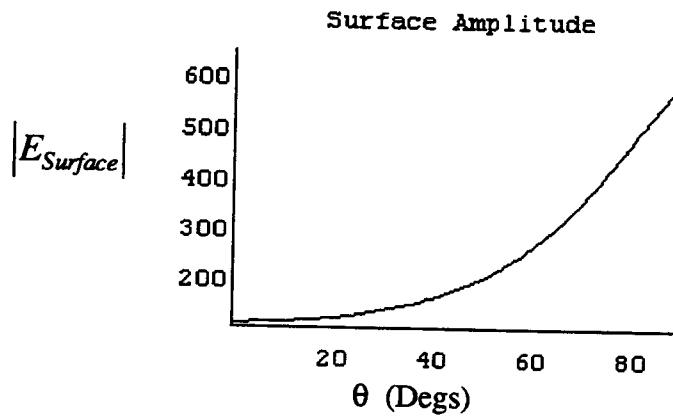


Figure 3.15: A plot of the magnitude of the surface contribution to generated electric field for S-polarization collinear optical mixing in dispersionless media.

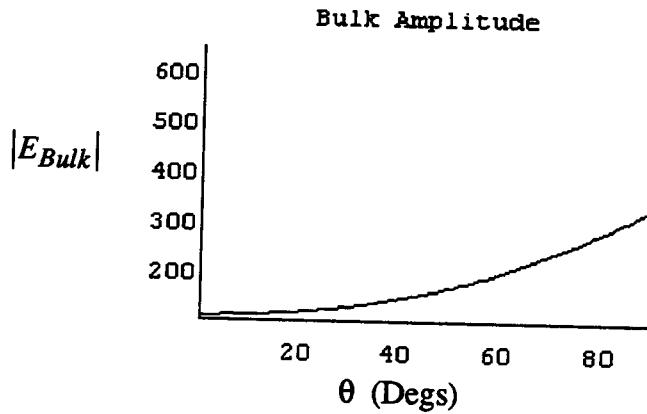


Figure 3.16: A plot of the magnitude of the bulk contribution to generated electric field for S-polarization collinear optical mixing in dispersionless media.

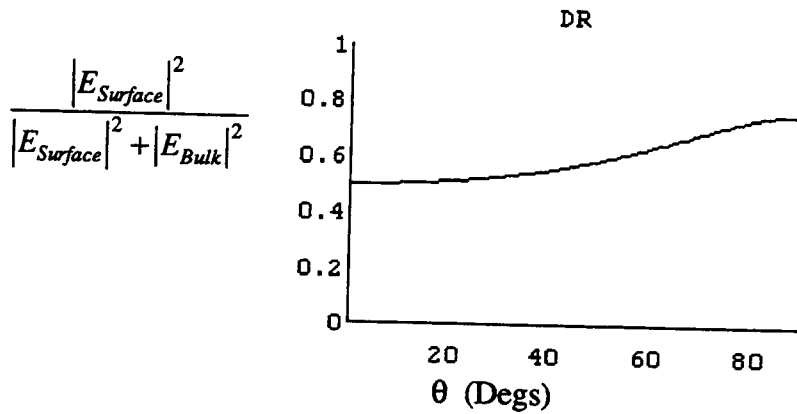


Figure 3.17: A plot of the DR versus incident angle for S-polarization collinear optical mixing in dispersionless media.

The overall angular dependence of the NLRCs can be understood by realizing that as the effective interaction length increases as the applied fields approach grazing incidence.

Viewing the plots, it is seen that the angular dependence is not strictly  $\sec \theta$ . Evidently, owing to linear refraction,  $\theta$  is smaller in the bulk median than in the surface layer, hence the bulk field is smaller. One can see the complicated  $\theta$  dependence that originates from the nonlinearly active layers (i.e.,  $\mathcal{E}_{BC}$ ).

The reader should be reminded that there is a question about the accuracy of the model at high angles. Since the (linear) surface-to-vacuum reflection coefficient is large at high angles, the single reflection approximation is suspect. At near-grazing angles, it is likely that there may be many reflections before the generated light escapes into the cover region; the whole structure might be construed as a leaky waveguide.

Very little was said about the phases,  $\tan \phi = \text{Im}[E] / \text{Re}[E]$ , of the bulk- and surface-generated  $S$ -waves for collinear optical mixing. As it happens, the phase difference,  $\phi_{\text{surface}} - \phi_{\text{bulk}}$ , is small and nearly constant over all  $\theta$ . Only if the dispersion is non-zero is there a significant phase difference, although it is still nearly constant. Note that thin-film interferences are not seen as  $d \ll \lambda$ .

A potentially useful  $S$ -wave optical mixing RISOM experiment is predicted for grazing angles of incidence. An experimental setup might appear like that shown in Figure 3.18. Figure 3.18 shows a harmonic generation experiment, but the same configuration may be used for any collinear OM experiment. Mixed waves are generated from the surface and bulk of the sample and sent to a detector. If at normal angles of incidence the bulk and surface signals are equal, at near-grazing angles the surface signal is somewhat enhanced. According to Figure 3.17, the surface signal is enhanced over the bulk by a factor of roughly 2 at 89 degrees. By observing the total mixed signal as a function of angle and fitting it to the RISOM expressions, the bulk and surface contributions may be separately deduced.

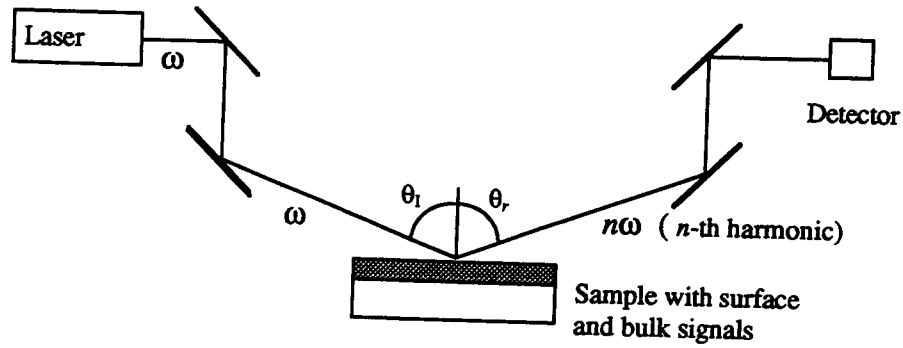


Figure 3.18: An S-wave optical mixing RISOM experiment performed at near-grazing angles.

### 3.6.1.2 S-polarization NLRC for Collinear $n$ HG in Media with Dispersion

So far the results presented for  $S$ -wave  $n$ HG are for perfect phase-matching (collinear geometry) processes, and since dispersion is not considered, the bulk and surface contributions to the observed intensity are maximum. The inclusion of dispersion will not change the angular dependencies nor destroy the perfect phase-matching. However, if the dispersions of the surface and bulk (medium) are different, the waves generated in the medium with the smallest dispersion will contribute more to the overall observed intensity. Since the index of refraction of the incident fields is different from the index for the generated wave,  $n_t^2 - n_s^2$  is non-zero and so is  $\Phi$  (see (3.47) and (3.48)). The generated intensity is, therefore, less than it would be for the dispersionless case. The angular dependencies, which were all contained in the  $\sec \theta$  and  $\mathcal{E}_{BC}$  factors, remain virtually unchanged. Suppose the surface has a smaller dispersion than the bulk linear index. Then the bulk-generated wave will be smaller due to a larger phase-mismatch and smaller  $(n_t^2 - n_s^2)^{-1}$  than the surface-generated wave. It is not likely that in a real situation, there would be any control over the dispersion of the bulk or surface. When doing experiments on a real sample, a separate measurement of the dispersion of surface and bulk would have to be made in order to determine to what extent the surface is coincidentally enhanced

over bulk. Since the experimenter may not have any control over the dispersions, separation of the contributions must be achieved by introducing additional phase-mismatch by use of non-collinear geometry.

### 3.6.1.3 *P*-polarization NLRC for Collinear $n$ HG in Dispersionless Media

Analysis of *P*-waves is more complicated than that of *S*-polarization. In the case of *P*-waves, the polarization is in the plane of incidence, hence there is a possible angle between the effective source wavevector and the polarization vector. In the development of the field amplitudes, this angle was labeled  $\alpha$ . (Refer to Figure 3.10 or Figure B.2 in Appendix B.) It is more useful to consider the angle  $\beta \equiv \pi - (\theta_s + \alpha)$ , which gives the direction of the polarization with respect to the upwards-going vacuum-surface interface normal. Again,  $\theta_s$  is the polar angle of the effective source wavevector as measured *inside* the medium. Recall in earlier discussions, there exists a NLBA for which the generated field is zero. This occurs at reflection angle (as measured in the vacuum) such that  $\theta_s = \beta - \pi$  for either nonlinear layer. (The reader should be aware that the field at the NLBA is never really zero; this is predicted in Ju's microscopic model.) Note that the value of the NLBA depends on the polarization direction as given by  $\beta$  rather than strictly a ratio of the linear indices of refraction as with the linear Brewster's angle. As a first trial, the polarization directions of the surface region, given by  $\beta_2$ , and the bulk,  $\beta_3$ , are considered equal. Note there can be no NLBA condition for the surface layer when  $\beta_2 < \pi/2 + \theta_{c2}$ , where  $\theta_{c2}$  is the critical angle at the vacuum-surface interface. For these values of  $\beta_2$ , the source wavevector can never be parallel to the polarization vector even if the applied fields are incident parallel to the surface. The same holds for the NLBA in the bulk median and angle  $\beta_3$ . Hence, there are four possible cases that can be investigated.

As illustrated in plots of the surface and bulk fields (Figure 3.19 and Figure 3.20) when  $\beta_2 = 105$  degrees  $< \pi/2 + \theta_{c2}$  and  $\beta_3 = \beta_2 < \pi/2 + \theta_{c3}$  neither layer experiences a NLBA condition. Note that the critical angles for the surface and bulk are  $\theta_{c2} = 42$  degrees and  $\theta_{c3} = 75$  degrees for the indices chosen. Another possibility is that a NLBA is

present for the surface ( $\beta_2 = 150$  degrees  $> \pi/2 + \theta_{c2}$ ), but not for the bulk median. This case is illustrated in Figure 3.21 and Figure 3.22. Only if the index of the surface is less than the index of the bulk median is this case relevant. Next, it is possible that  $\beta_2 = \beta_3$  is large enough that both layers have a NLBA. Without dispersion, the NLBAs of the two layers are the same as shown in Figure 3.23 for the surface field and Figure 3.24 for the bulk. Here  $\beta_2$  and  $\beta_3$  are 170 degrees such that  $\beta_3 = \beta_2 > \pi/2 + \theta_{c3} > \pi/2 + \theta_{c2}$ . None of the cases shown here are useful for RISOM experiments. For instance, the DR in the case when both layers have a NLBA (Figure 3.25) indicates that the surface signal cannot be more than twice the bulk signal if at normal incidence they are equal. This is no better than the case for the grazing incidence experiment for *S*-waves. Despite these discouraging results, some credence comes to these calculations in that the *S*-wave and *P*-wave values are equal at  $\theta = 0$ . The  $\sec \theta$  dependence is again apparent in the field amplitudes at high angles. Also, although not shown here is the expected phase change at the location of the NLBA.

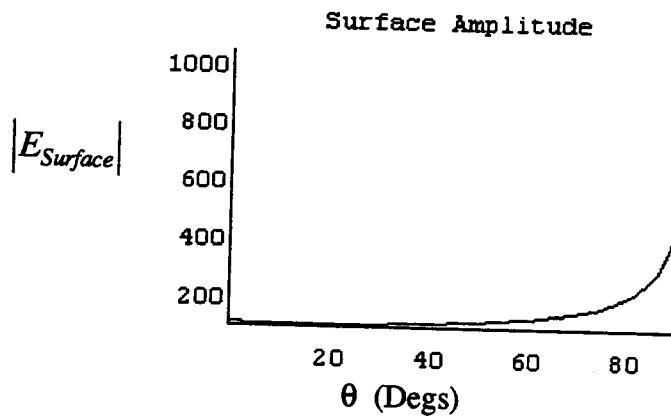


Figure 3.19: The *P*-wave surface amplitude versus incident angle for collinear *n*HG in dispersionless media ( $\beta = 105$  degrees).

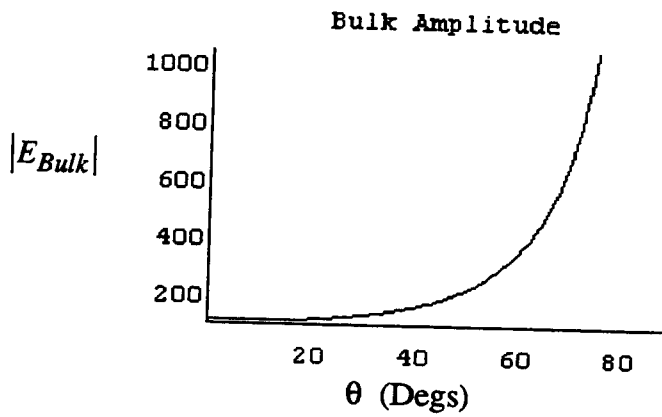


Figure 3.20: The  $P$ -wave bulk amplitude versus incident angle for collinear  $n$ HG in dispersionless media ( $\beta = 105$  degrees).

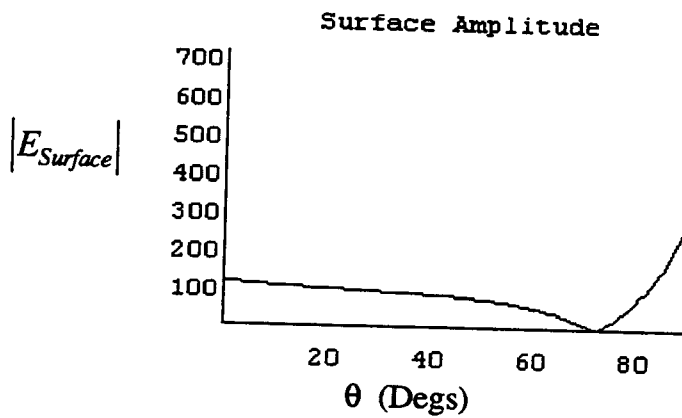


Figure 3.21: The  $P$ -wave surface amplitude versus incident angle for collinear  $n$ HG in dispersionless media ( $\beta = 150$  degrees).



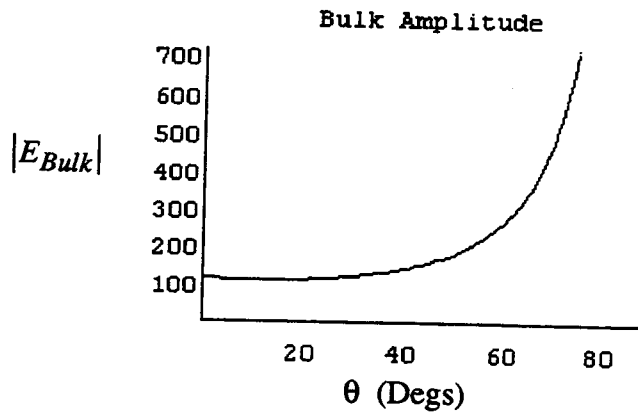


Figure 3.22: The  $P$ -wave bulk amplitude versus incident angle for collinear  $nHG$  in dispersionless media ( $\beta = 150$  degrees).

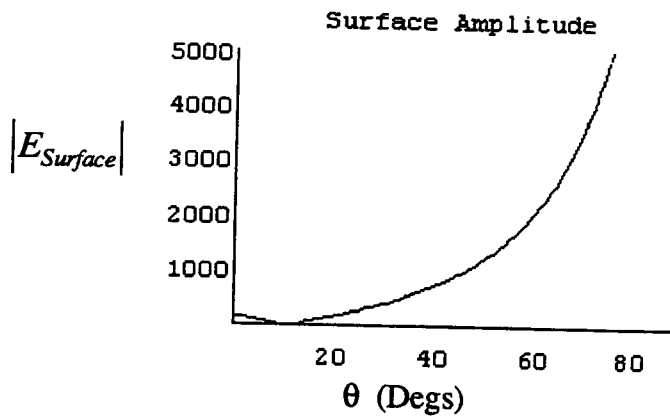


Figure 3.23: The  $P$ -wave surface amplitude versus incident angle for collinear  $nHG$  in dispersionless media ( $\beta = 170$  degrees).

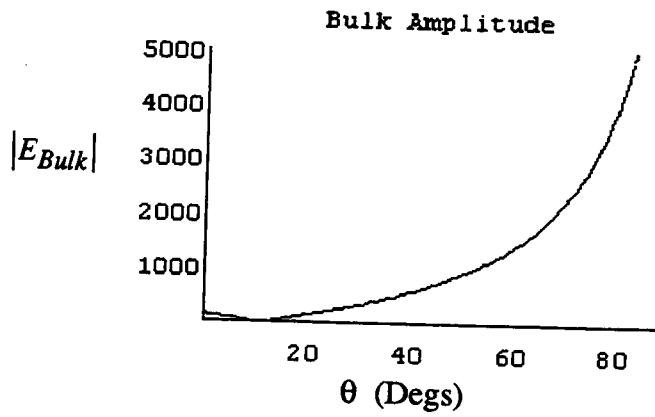


Figure 3.24: The *P*-wave bulk amplitude versus incident angle for collinear *n*HG in dispersionless media ( $\beta = 170$  degrees).

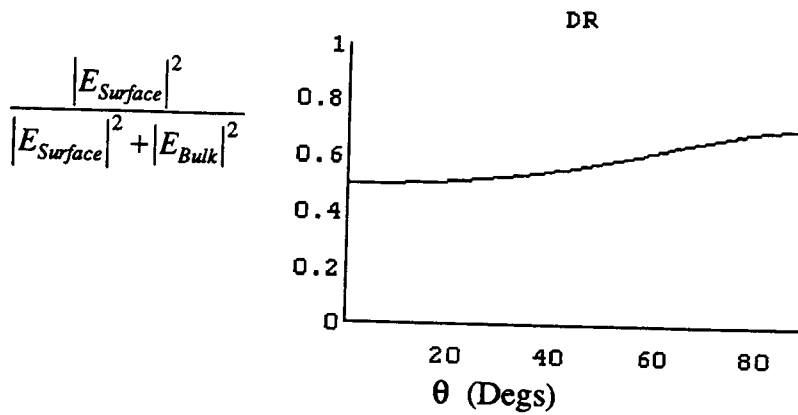


Figure 3.25: A plot of the DR versus incident angle for *P*-polarization collinear optical mixing in dispersionless media ( $\beta = 170$  degrees).

The final possibility the experimenter may encounter occurs when the linear index of the surface is larger than the bulk. This is possible on a  $\text{TiO}_2$  bulk crystal where  $\text{Ti}_2\text{O}_3$  comprises a surface layer. If  $\beta_3 > \pi/2 + \theta_{c3}$  and one is lucky enough to have no NLBA in the surface, the bulk signal would disappear at its NLBA leaving only the surface signal.

### 3.6.2 P-Wave Collinear Geometry SHG With Dispersion: Further Analysis of the NLBA

With the inclusion of dispersion a phase-mismatch ensues. From the discussion on the NLRCs for *S*-polarization with dispersion, it is expected that the overall intensity is strongly diminished by the  $(n_t^2 - n_s^2)^{-1}$  factor. In addition, as (positive) dispersion is included the NLBA decreases. This is evident by realizing that  $n_s \sin \theta_s = n_t \sin \theta_t$ , where  $n_s$  depends on  $n(\omega)$  (see (3.5)) and  $n_t = n(2\omega)$  for SHG. If the dispersion of the bulk and surface are different the NLBAs will differ. This offers another way to emphasize the surface contribution.

In order to study the effects of dispersion on the angular dependence of collinear optical harmonic generation, a small frequency dependence for the linear indices of refraction was introduced. The dispersion was chosen to be very small, 0.001 for the surface and 0.002 for the bulk. Hence, the index for the surface at the fundamental (532 nm) was 1.500, but 1.501 for the second harmonic. The bulk index was 1.55 for the fundamental and 1.552 for the second harmonic. Typical dispersions are 5 to 10 times larger than this for dielectric solids. To emphasize the sensitivity of the dispersion to the NLBA, it was decided to use these small dispersions.

The surface signal as a function of incident angle for collinear SHG for  $\beta = 150$  degrees is shown in Figure 3.26. The bulk signal's angular dependence is plotted in Figure 3.27. The difference between surface and bulk Brewster's angles is 7.3 degrees. Utilizing this difference, separation of surface and bulk can be favourable as shown in the DR in

Figure 3.28. When the angle of incidence is chosen to coincide with the NLBA of the bulk, 46 degrees, the surface signal is small, but finite. Also evident are the changes in phase as evidenced by the CR in Figure 3.29.

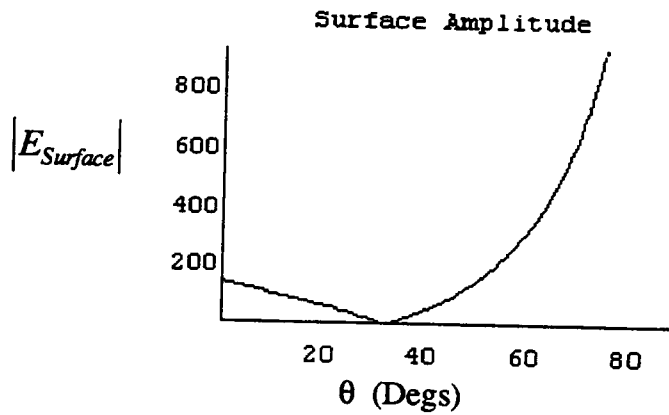


Figure 3.26: The *P*-wave surface contribution for collinear SHG ( $\beta = 150$  degrees), with 0.001 dispersion.

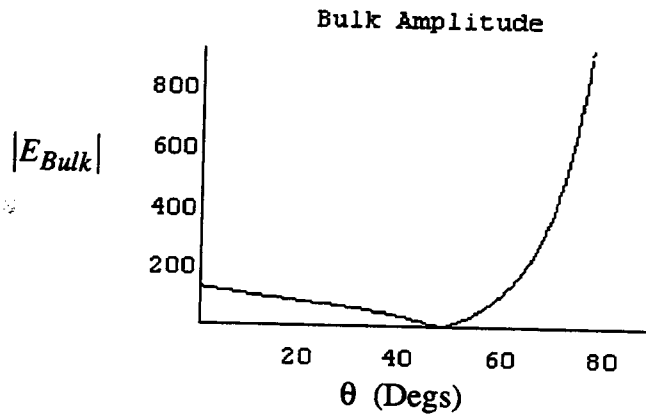


Figure 3.27: The *P*-wave bulk contribution for collinear SHG ( $\beta = 150$  degrees), with 0.002 dispersion.

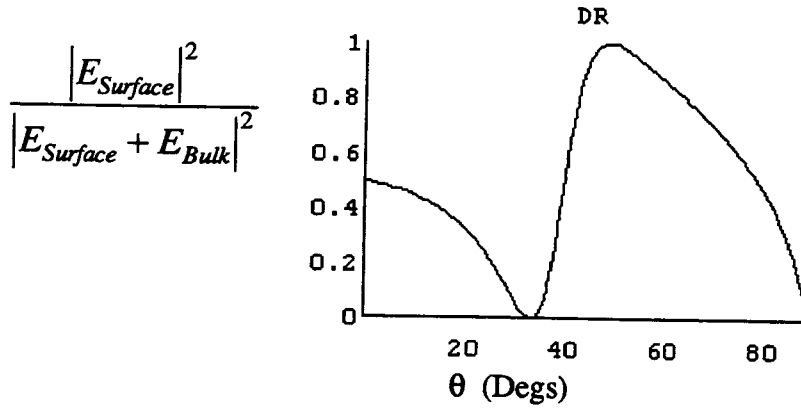


Figure 3.28: The *P*-wave DR for collinear SHG ( $\beta = 150$  degrees), with dispersion.

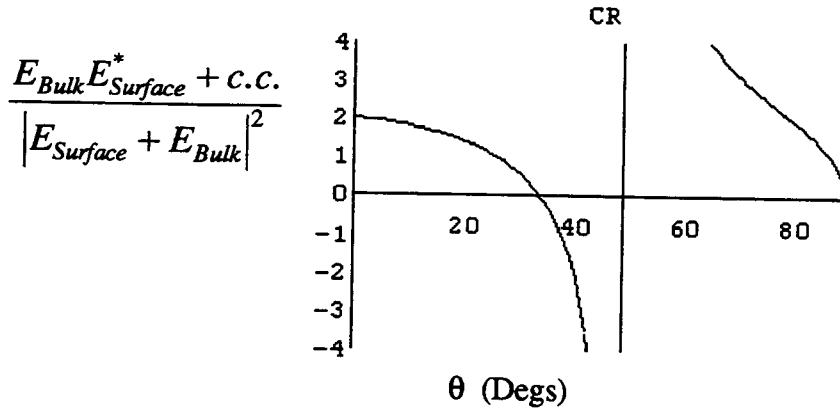


Figure 3.29: The *P*-wave CR versus incident angle for collinear SHG ( $\beta = 150$  degrees), with dispersion.

Since the bulk signal is usually much larger, the experimenter would apply collinear *P*-waves to a sample and find an incident angle that extinguishes the majority of the

generated waves. This will be the NLBA of the bulk. The remaining signal is generated in the surface layer. Changing the nature of the surface should result in an obvious change in the remaining signal from the surface. For instance, the surface resonance yielding a strong SHG signal might fall off-resonance if the surface is altered by thermal desorption.

### 3.7 Analysis of RISOM in Non-Collinear Geometry

After a brief review of non-collinear geometry experiments, done in Section 3.7.1, an experiment that introduces a mismatch via breaking the degeneracy in polar angles is discussed. Initially, the analysis is simplified by considering only SHG in non-dispersive media. This is presented in Section 3.7.1.1. In Section 3.7.1.2, an experiment which uses the azimuthal angles to achieve phase-mismatched non-collinear SHG is examined. Then non-collinear phase-mismatched SHG with a NLBA is investigated. Harmonic generation in media with dispersion is briefly discussed. The remainder of the analysis investigates non-collinear geometry of three-wave and four-wave mixing. Section 3.7.2 considers TWM for media with dispersion. Finally, an analysis of FWM is done in Section 3.7.3 for media with dispersion.

#### 3.7.1 Non-Collinear Geometry SHG Without Dispersion

There is additional angular dependence of the nonlinear reflection coefficients arising from phase-mismatched situations. This dependence occurs in the factor  $\mathcal{E}_{BC}$ . Dispersion, a parameter the experimenter has little choice of, is one way mismatched conditions arise. Another way to introduce phase-mismatch is to slightly change the direction of one (or more) of the incident wavevectors from the phase-matched case. For harmonic generation a mismatch is introduced by going to a non-collinear geometry. Notice (from (3.5)) the vector sum of the incident wavevectors is the effective source wavevector,  $k_s$ . The generated wave has a wavevector magnitude  $k_t = n(2\omega) (2\omega)/c$ , where  $\omega$  is the fundamental frequency and a direction such that  $\vec{k}_t - \vec{k}_s \equiv \Delta\vec{k}$  is minimized. For SHG  $\vec{k}_t$

comes out exactly half-way in between the two incident waves. For instance: if  $\theta_{1,I} = 44$  degrees,  $\theta_{2,I} = 46$  degrees, and  $\phi_{1,I} = \phi_{2,I} = 0$  degrees; then  $\theta_{3,I} = |\theta_{2,I} + \theta_{1,I}|/2 = 45$  degrees and  $\phi_{3,I} = 0$  degrees.

Observing the forms of  $\mathcal{E}_{BC,2}$  and  $\mathcal{E}_{BC,3}$ , it can be deduced that when there is a phase-mismatch the angular dependence on  $\theta$  are slightly different for the surface and the bulk median. This is only true because the linear indices of refraction are dissimilar. This difference, as will be illustrated in the results to follow, comes from two terms in  $\mathcal{E}_{BC,2}$  involving  $(\sin \phi_{s2} - \sin \phi_{i2})$  and  $(\sin \phi_{s3} - \sin \phi_{i3})$  in  $\mathcal{E}_{BC,3}$ . This difference is further increased by dispersion, and exacerbated when the dispersion of surface and bulk are different. It is noted that when dispersion is included, it is possible to mismatch the surface-generated wave and greatly mismatch the bulk wave due to this difference. The experimenter may thereby generate a weak surface signal and a minuscule bulk signal. This is due to the enormous sensitivity of the  $(n_t^2 - n_s^2)^{-1}$  factor to the phase-mismatch.

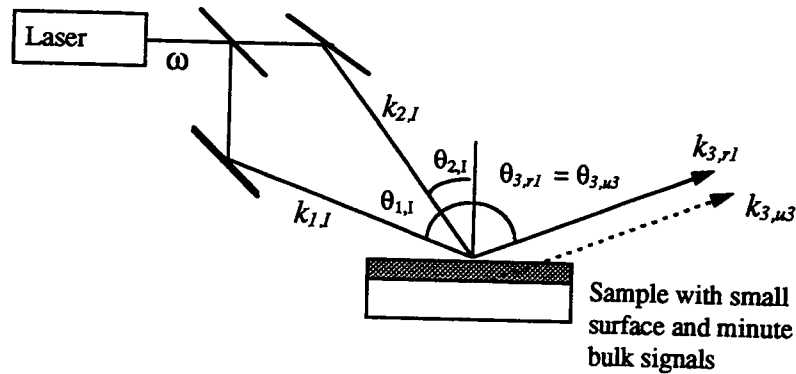


Figure 3.30: A non-collinear SHG experiment utilizing only the polar angles to introduce phase-mismatch.

Non-collinear experiments for SHG might be designed as shown in Figure 3.30. Here the incident frequency source is split into two waves,  $k_{1,I}$  and  $k_{2,I}$ , which are incident on

the sample with slightly different (polar) angles of incidence,  $\theta_{1,I}$  and  $\theta_{2,I}$ . The azimuthal angles,  $\phi_{1,I}$  and  $\phi_{2,I}$ , are zero. Let the difference between these incidence angles be  $\delta\theta_{1,2} \equiv |\theta_{1,I} - \theta_{2,I}|$ . The second harmonic 'reflects' off the surface with wavevector,  $k_{3,r,I}$ , at angle  $\theta_{3,r,I}$ . It is high in intensity because the phase-mismatch in the surface has been selected (by choice of  $\theta_{1,I}$  and  $\theta_{2,I}$ ) to be small. The bulk signal necessarily 'reflects' at angle  $\theta_{3,\mu,3} = \theta_{3,r,I}$ . Note that when there is dispersion, these angles may not be equal.

Another way to design non-collinear experiments for SHG is shown in a perspective view in Figure 3.31. The sample is shown face up, tilted somewhat in this diagram to show all the directions of the waves. Again, the incident frequency source is split into two waves,  $k_{1,I}$  and  $k_{2,I}$ . These waves are incident on the sample with the same (polar) angle of incidence,  $\theta_{1,I} = \theta_{2,I}$ . However, the azimuthal angles,  $\phi_{1,I}$  and  $\phi_{2,I}$ , are slightly different. (They are exaggerated in Figure 3.31.) Let the difference between these angles be  $\delta\phi_{1,2} \equiv |\phi_{1,I} - \phi_{2,I}|$ . The second harmonic 'reflects' off the surface with wavevector,  $k_{3,r,I}$ , at angles  $\theta_{3,r,I}$  and  $\phi_{3,r,I} = (\phi_{1,I} + \phi_{2,I})/2$ . For reference, the definition  $\phi_{3,\mu,3} \equiv \phi_{3,r,I} \equiv 0$

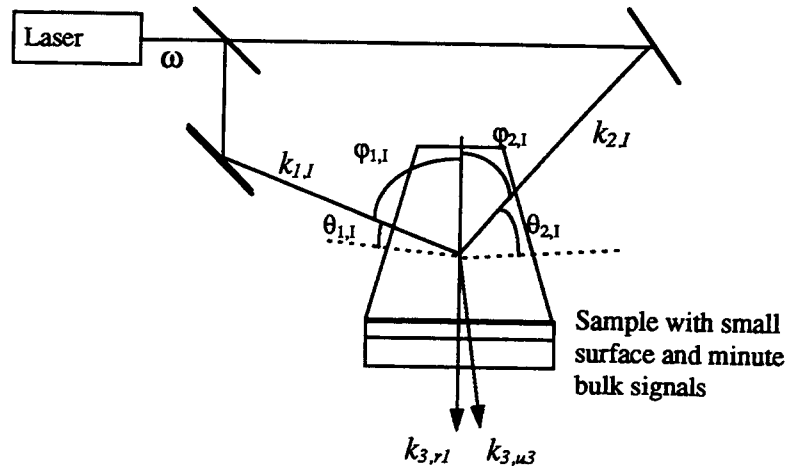


Figure 3.31: A non-collinear SHG experiment utilizing only the azimuthal angles to introduce phase-mismatch.



is made, thus  $\phi_{1,I} = -\phi_{2,I}$ . When the phase-mismatch in the surface is small (done by choosing  $\phi_{1,I}$  and  $\phi_{2,I}$ ), the surface-generated wave is high in intensity. The bulk signal 'reflects' at angles  $\theta_{3,\mu 3} = \theta_{3,rI}$ , and  $\phi_{3,\mu 3} = \phi_{3,rI}$ . As before, when there is dispersion, these angles may not be equal.

### 3.7.1.1 Polar Angle Phase-Mismatched S-Polarization SHG Without Dispersion

The analysis of the RISOM expressions for non-collinear SHG begins with the  $\delta\theta_{1,2} \neq 0$ ,  $\delta\phi_{1,2} = 0$  experiment (See Figure 3.30). Numerical calculations involving the RISOM expressions as a function of  $\delta\theta_{1,2}$  were made for several values of the harmonic 'reflected' angle,  $\theta_{3,rI} = \theta_{3,\mu 3}$ . In the calculations that follow, the angle of 'reflection',  $\theta_{3,rI}$ , is chosen. Values chosen are 15, 45, 60, and 89 degrees. Then the difference of the incidence angles,  $\delta\theta_{1,2}$ , is scanned. From  $\theta_{3,rI} = |\theta_{2,I} + \theta_{1,I}|/2$  and  $\delta\theta_{1,2} \equiv |\theta_{1,I} - \theta_{2,I}|$ , the angles of incidence (with respect to the normal),  $\theta_{1,I}$  and  $\theta_{2,I}$ , are found. The field amplitudes and related ratios can then be calculated and plotted versus  $\delta\theta_{1,2}$ .

The surface contribution to the generated field is shown in Figure 3.32 for  $\theta_{3,rI} = 15$  degrees as a function of  $\delta\theta_{1,2}$ . As expected, the value of the contribution at the peak, where  $\delta\theta_{1,2} = 0$ , is the value shown in Figure 3.15 at 15 degrees. This is evidence of the  $\mathcal{E}_{BC,2}$  factor's  $\theta_{3,rI}$  dependence and the  $\sec \theta_{3,rI}$  factor (referring to (3.47)). The  $(n_{t2}^2 - n_{s2}^2)^{-1}$  dependence as an overall  $(\delta\theta_{1,2})^{-2}$ -like dependence. Recall that the field is finite at  $\delta\theta_{1,2} = 0$ , since the limit of  $\mathcal{E}_{BC,2} \times (n_{t2}^2 - n_{s2}^2)^{-1}$  as  $n_{s2} \rightarrow n_{t2}$  is finite. Also evident is the  $\text{sinc}\Phi$  behaviour on the surface field amplitude. Surface contributions to the generated fields for the other choices of  $\theta_{3,rI}$  are shown in Figure 3.33 ( $\theta_{3,rI} = 45$  degrees), Figure 3.34 ( $\theta_{3,rI} = 60$  degrees), and Figure 3.35 ( $\theta_{3,rI} = 89$  degrees). Note that in Figure 3.35 the  $\text{sinc}\Phi$  has been suppressed by plotting only the maxima of the oscillation apparent in the other figures. The effect is to observe only the effect of  $\mathcal{E}_{BC,2} \times (n_{t2}^2 - n_{s2}^2)^{-1}$ . Comparing these plots the  $\mathcal{E}_{BC,2}$  dependence on  $\theta_{3,rI}$  for mismatched processes can be observed. As  $\theta_{3,rI}$  increases the dependence on the phase-mismatch (i.e.,  $\delta\theta_{1,2}$ ) becomes smaller. Hence, the plots appear broader with increasing  $\theta_{3,rI}$ .

Correspondingly, the bulk contributions to the generated fields show a similar dependence. Plots of the bulk field amplitude for the choices of  $\theta_{3,\mu 3} = \theta_{3,r,l}$  are shown in Figure 3.36 ( $\theta_{3,r,l} = 15$  degrees), Figure 3.37 ( $\theta_{3,r,l} = 45$  degrees), Figure 3.38 ( $\theta_{3,r,l} = 60$  degrees), and Figure 3.39 ( $\theta_{3,r,l} = 89$  degrees). As before, the peak intensities are consistent with Figure 3.16. The  $\theta_{3,r,l}$  dependence on  $\mathcal{E}_{BC,3}$  causes the bulk contribution to be less sensitive to  $\delta\theta_{1,2}$  than the surface -- the bulk plots for larger  $\theta_{3,r,l}$  are broader. Again, this is due to terms in the  $\mathcal{E}_{BCS}$  like  $(\sin \phi_{s2} - \sin \phi_{l2})$  and  $(\sin \phi_{s3} - \sin \phi_{l3})$ .

An example of a successful RISOM experiment is shown in the last plot in this series. Figure 3.40 shows a plot of the *total signal intensity* versus  $\delta\theta_{1,2}$  at  $\theta_{3,r,l} = 89$  degrees. Perhaps the most striking element of this experiment is the separation of the bulk and surface signals via phase-mismatch. The combination of these two signals gives an intensity profile in the shape of the Kaiser's Helmet; low dome-shaped sides with a sharp spike at the center. The domed part represents only bulk signal; the bulk signal is much less sensitive to mismatch, that is, less sensitive to changes in  $\delta\theta_{1,2}$ . The bulk signal at larger difference angles  $\delta\theta_{1,2}$  is attenuated by  $10^3$  to  $10^5$  over its maximum signal. The sharp spike represents the surface signal atop the bulk signal. The surface signal is approximately five times more sensitive to  $\delta\theta_{1,2}$  than the bulk signal at  $\delta\theta_{1,2}$  at  $\theta_{3,r,l} = 89$  degrees. Hence at  $\delta\theta_{1,2} = 0.1$  degrees, the surface-to-bulk ratio is  $1.6 \times 10^{-2}$ , it is in favour of the bulk signal. At  $\delta\theta_{1,2} = 0$ , however, the surface-to-bulk ratio leaps up to 2.3. A large  $\theta_{3,r,l}$  was chosen for the same reason as the first experiment -- at grazing angles of incidence, the surface signal is roughly twice that of the bulk signal (if the bulk signal and surface signal are equal at normal incidence).

Any actual experiment would not use plane waves as the model has assumed. The incident light would be in the form of beams, Gaussian in nature. Furthermore, these beams would be focused onto the sample as to maximize the induced polarization, and

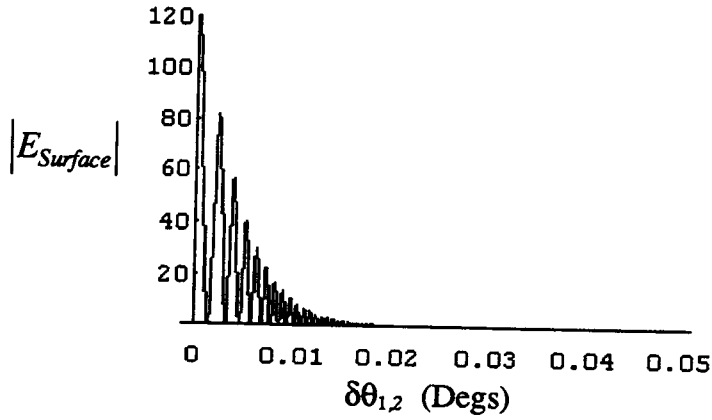


Figure 3.32: Plot of the surface contribution of the field versus  $\delta\theta_{1,2}$ . The process is non-collinear SHG with  $\theta_{3,r1} = 15$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

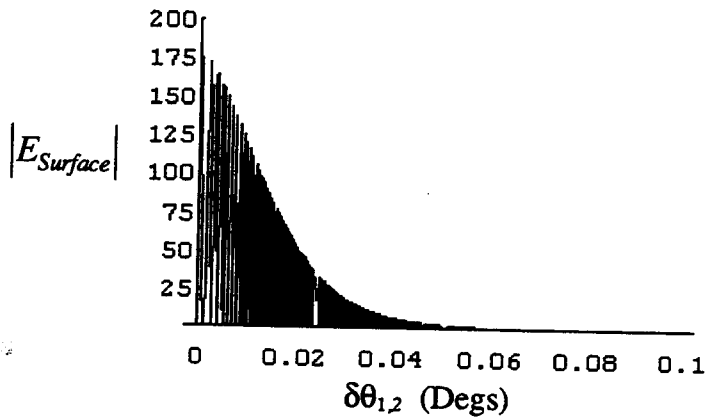


Figure 3.33: Plot of the surface contribution of the field versus  $\delta\theta_{1,2}$ . The process is non-collinear SHG with  $\theta_{3,r1} = 45$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

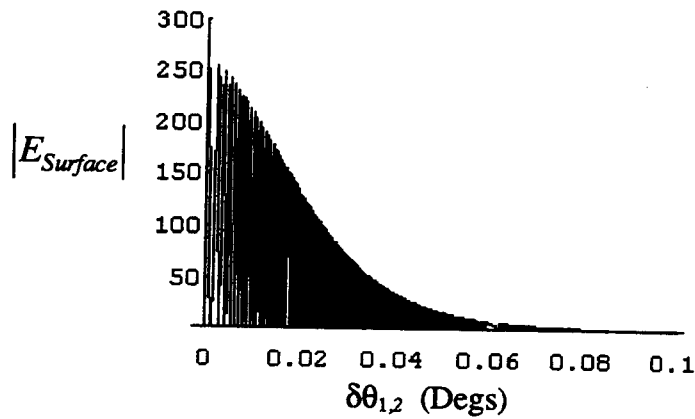


Figure 3.34: Plot of the surface contribution of the field versus  $\delta\theta_{1,2}$ . The process is non-collinear SHG with  $\theta_{3,r,l} = 60$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

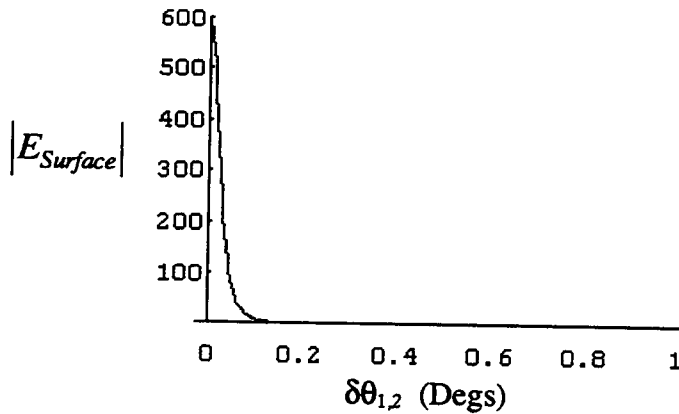


Figure 3.35: Plot of the surface contribution of the field versus  $\delta\theta_{1,2}$  (*sinc* modulation suppressed). The process is non-collinear SHG with  $\theta_{3,r,l} = 89$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

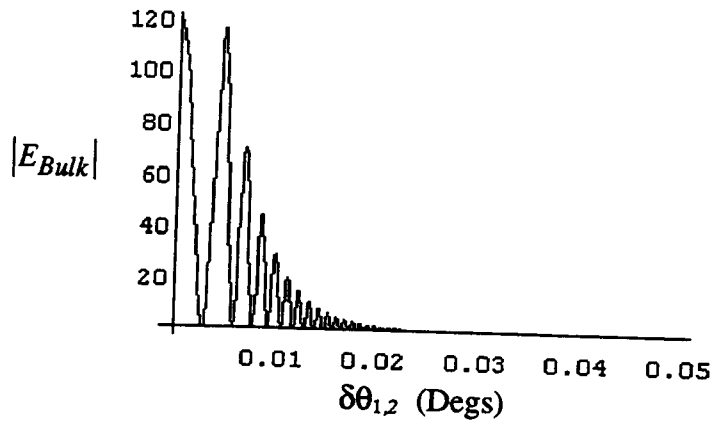


Figure 3.36: Plot of the bulk contribution of the field versus  $\delta\theta_{1,2}$ . The process is non-collinear SHG with  $\theta_{3,r,l} = 15$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

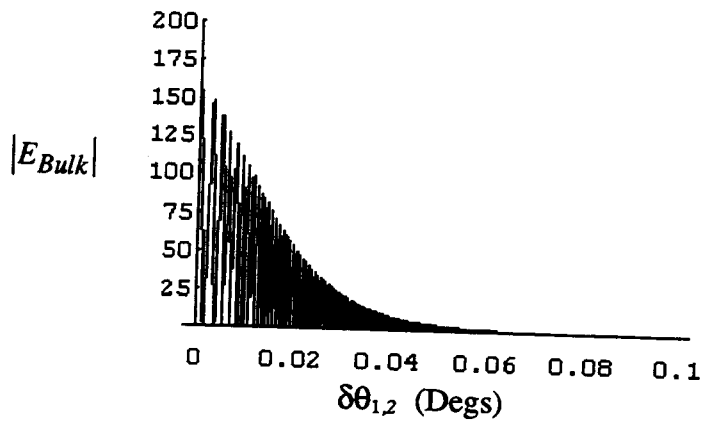


Figure 3.37: Plot of the bulk contribution of the field versus  $\delta\theta_{1,2}$ . The process is non-collinear SHG with  $\theta_{3,r,l} = 45$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

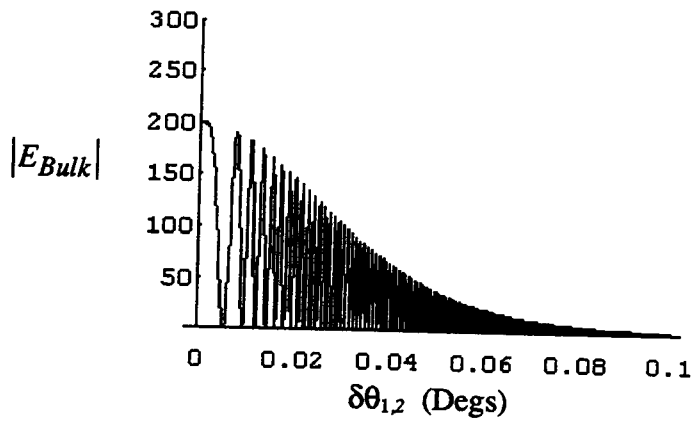


Figure 3.38: Plot of the bulk contribution of the field versus  $\delta\theta_{1,2}$ . The process is non-collinear SHG with  $\theta_{3,r1} = 60$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

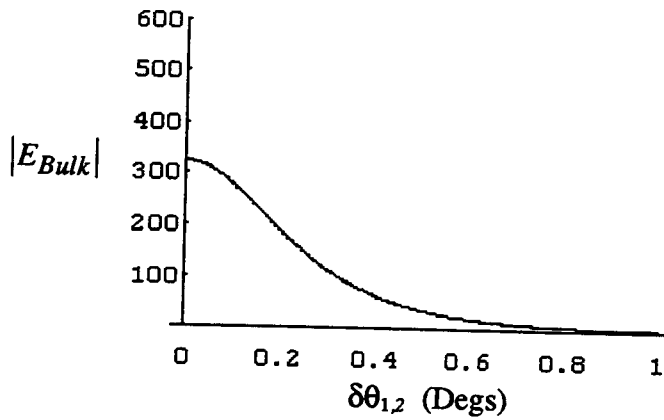


Figure 3.39: Plot of the bulk contribution of the field versus  $\delta\theta_{1,2}$  (*sinc* modulation suppressed). The process is non-collinear SHG with  $\theta_{3,r1} = 89$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

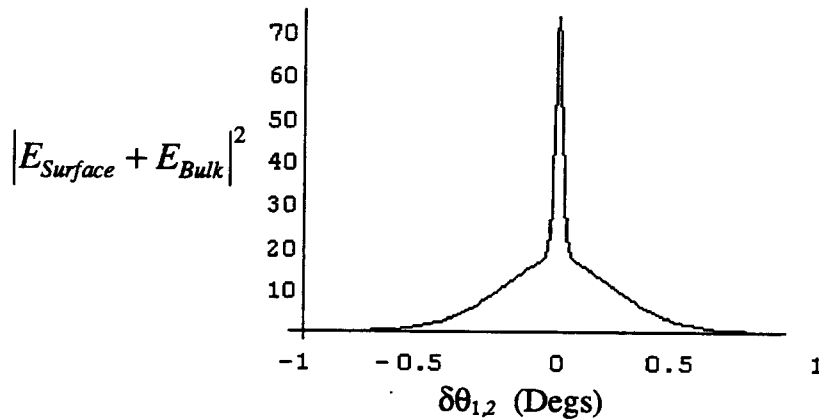


Figure 3.40: A plot of the ratio of total intensity (x 6000) versus  $\delta\theta_{1,2}$  (*sinc* modulation suppressed). The process is non-collinear SHG with  $\theta_{3,rl} = 89$  degrees,  $\delta\phi_{1,2} = 0$ , no dispersion.

hence, the harmonic signal. Decomposing the focused Gaussian beam into seven plane waves, and repeating the calculation for each plane wave leads to a more experimentally realistic prediction. Estimates were made for a single Gaussian beam that is focused to a 10 micron spot (the interaction region) using a 1 meter focal length lens. The beam contains a continuum (approximated by seven waves) of incident waves that mix at the interaction region which is partially in the surface and partially in the bulk region. Each plane wave in the beam, except the central one, mix together with a non-zero mismatch. The center of the beam is incident at 89 degrees. Due to the different sensitivities of phase-mismatch, the SHG beam has an angular profile. The angular profile looks like a broadened version of the Kaiser's Helmet, Figure 3.40. However, it is 18 times broader. The intensity at the center of the SHG beam is due to both surface and bulk. The angular width of this center portion (the beam is a cone) is about 0.02 degrees (full-width half-

maximum or one-fifth of the full beam-width of 1 degree). The outer portion of the SHG beam has a much lower intensity and is due primarily to light generated from the bulk layer. The signal at the center portion of the beam yields a surface-to-bulk signal ratio of (approximately) 4.

### 3.7.1.2 Azimuthal Angle Phase-Mismatched S-Polarization SHG Without Dispersion

Non-collinear SHG using the azimuthal angles, the experiment represented in Figure 3.31, is now investigated. Again, the incident frequency source is split into two waves,  $k_{1,I}$  and  $k_{2,I}$ . However, now these waves are incident on the sample with the same angle of incidence,  $\theta_{1,I} = \theta_{2,I}$ , but differing azimuthal angles,  $\phi_{1,I}$  and  $\phi_{2,I}$ . The difference between these angles was notated as  $\delta\phi_{1,2} \equiv |\phi_{1,I} - \phi_{2,I}|$ . The second harmonic 'reflection' occurs at angles  $\theta_{3,rI} = \theta_{3,\mu3} = \theta_{1,I} = \theta_{2,I}$  and  $\phi_{3,rI} = (\phi_{1,I} + \phi_{2,I})/2$ . If the azimuthal angles  $\phi_{3,\mu3} = \phi_{3,rI}$  are defined to be zero, then  $\phi_{1,I} = -\phi_{2,I}$ . The (polar) angles of incidence,  $\theta_{1,I} = \theta_{2,I}$ , provide the  $\mathcal{E}_{BC}$  overall dependence to the signal and bulk signals. The fields change with  $\delta\phi_{1,2}$  due to both the  $\mathcal{E}_{BC} \times (n_i^2 - n_s^2)^{-1}$  factor and the *sinc*  $\Phi$ . From a RISOM perspective it is unfortunate that, with respect to  $\delta\phi_{1,2}$ ,  $\mathcal{E}_{BC,2}$  effects the surface and bulk median fields very little.

It was found that the bulk and surface waves have essentially the same  $\delta\phi_{1,2}$  dependence. Moreover, the  $\delta\phi_{1,2}$  dependence remains virtually constant for all  $\theta_{3,rI}$ . As an example, see Figure 3.41 and Figure 3.42. For  $\theta_{3,rI} = 60$  degrees the surface and bulk signals are nearly identical. Plots for other  $\theta_{3,rI}$  are not shown.

### 3.7.2 Non-Collinear Geometry TWM With Dispersion

For TWM processes in media with dispersion a non-collinear geometry must be used in order to phase-match. When the surface and bulk have different dispersions there is a finite angular separation between the surface and bulk waves. The generated waves



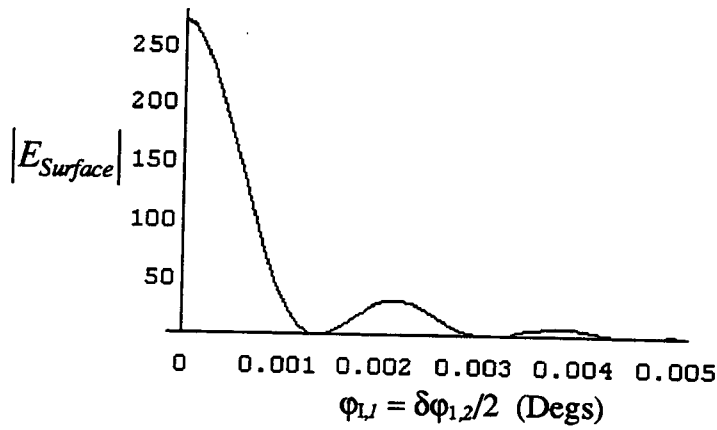


Figure 3.41: A plot of the surface field amplitude versus  $\phi_{LI}$ . The process is non-collinear SHG with  $\theta_{3,r1} = \theta_{3,\mu3} = 60$  degrees, no dispersion.

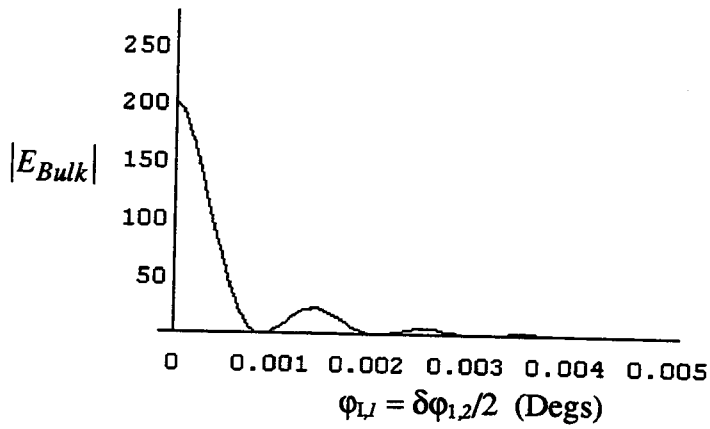


Figure 3.42: Plot of the bulk field amplitude versus  $\phi_{LI}$ . The process is non-collinear SHG with  $\theta_{3,r1} = \theta_{3,\mu3} = 60$  degrees, no dispersion.

‘reflect’ at the angle for which the phase-mismatch is minimum, given the incident wavevectors and differences in dispersion. This is illustrated in the expression for the set of surface reflection angles,

$$\cos\theta_{3,r1} = \frac{\omega_1 n_2(\omega_1) \cos\theta_{1,l1} + \omega_2 n_2(\omega_2) \cos\theta_{2,l1}}{(\omega_1 + \omega_2) n_2(\omega_1 + \omega_2)} \quad (3.49)$$

and

$$\cos\phi_{3,r1} = \frac{\omega_1 n_2(\omega_1) \sin\theta_{1,l1} \cos\phi_{1,l1} + \omega_2 n_2(\omega_2) \sin\theta_{2,l1} \cos\phi_{2,l1}}{(\omega_1 \cos\theta_{3,r1} + \omega_2 \cos\theta_{3,r1}) n_2(\omega_1 + \omega_2)}, \quad (3.50)$$

and the bulk reflected angles,

$$\cos\theta_{3,\mu3} = \frac{\omega_1 n_3(\omega_1) \cos\theta_{1,l1} + \omega_2 n_3(\omega_2) \cos\theta_{2,l1}}{(\omega_1 + \omega_2) n_3(\omega_1 + \omega_2)} \quad (3.51)$$

and

$$\cos\phi_{3,\mu3} = \frac{\omega_1 n_3(\omega_1) \sin\theta_{1,l1} \cos\phi_{1,l1} + \omega_2 n_3(\omega_2) \sin\theta_{2,l1} \cos\phi_{2,l1}}{(\omega_1 \cos\theta_{3,\mu3} + \omega_2 \cos\theta_{3,\mu3}) n_3(\omega_3)}. \quad (3.52)$$

The angular separation between surface and bulk waves for a SFG process was calculated using plane waves. In this example a 500 nm wave is incident at  $\theta_{1,l} = 88.9$  degrees and a 600 nm wave incident at  $\theta_{2,l} = 89.87$  ( $\delta\theta_{1,2} = .97$  degrees). The dispersion for the surface is 0.001 and is 0.002 for the bulk. The surface is minimally phase-mismatched such that the generated ‘sum frequency’ wave is emitted at  $\theta_{3,r1} = 89.33$  degrees, whereas the bulk signal phase-matches at  $\theta_{3,\mu3} = 89.68$  degrees. When performing this SFG experiment, the angular discrimination between surface and bulk waves is  $|\theta_{3,r1} - \theta_{3,\mu3}| = .35$  degrees.

Calculations were made for two Gaussian beams, one at frequency 500nm and one at 600 nm. Each Gaussian beam was focused to a 10 micron spot (the interaction region) using a 1 meter lens. As before, each Gaussian beam was decomposed into seven plane waves. The center of the mixed beam ‘reflects’ at 89.55, degrees as in the previous example with plane waves. The center of the incident beams were identical to those used in that example ( $\theta_{1,I} = 88.9$  degrees and  $\theta_{2,I} = 89.87$ ). The mixed beam shows an angular profile which has two distinct parts: a bright spot surrounded by a much larger dim halo. The angular profile shows a bright spot that is off center from the center of the halo. This angular profile is shown in Figure 3.43. The sharp spike represents the location of the bright spot; intensity generated primarily from the surface. The halo is shown in the profile as a wide, gently sloping peak that is located at 89.82 degrees. The ‘zero’ in Figure 3.43 is actually located at 89.50 degrees. The entire angular width of the beam is 1.6 degrees. This experiment yields a surface-to-bulk signal ratio of 3.4, when observing only the light from the bright spot.

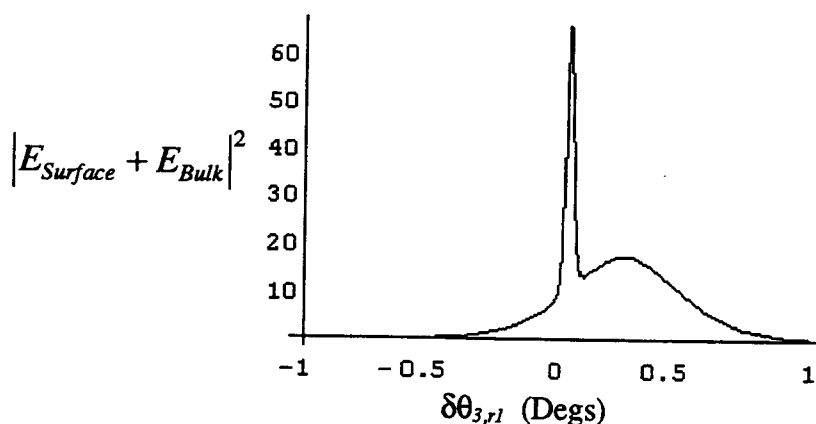


Figure 3.43: A plot of the polar angular intensity profile versus ‘reflection’ angle. The process is Gaussian beam SFG with dispersion.

### 3.7.3 Non-Collinear Geometry FWM With Dispersion

As was seen in TWM in media with dispersion, the mixed bulk and surface waves have a finite angular separation. FWM processes may require different azimuthal angles as well as polar angles to properly phase-match. Hence, the angular discrimination for FWM is larger. In general, the angular discrimination improves for higher order processes. It is useful to calculate the angular separation between *plane* waves in this non-collinear geometry. This will help to understand the experimental situation when Gaussian beams are employed. Even for small dispersions (0.001 for the surface, 0.002 for the bulk), the separation is experimentally noticeable. Consider a CARS-like four-wave mixing experiment. In this example, wavelengths of 500 nm incident at  $\theta_{1,I} = 88.90$  degrees, 600 nm incident at  $\theta_{2,I} = 88.96$  degrees, and 550 nm incident at  $\theta_{3,I} = 88.91$  degrees are mixed. The result are two waves at  $\omega_4 = \omega_1 - \omega_2 + \omega_3$  (465 nm) ‘reflecting’ at angles  $\theta_{4,rI} = \theta_{4,\mu3}$ . All azimuthal angles were chosen to be zero. This choice of angles causes the surface-generated wave to be phase-matched at  $\theta_{4,rI} = 89.74$  degrees. In order to achieve phase-matching in the bulk, the incident angles must be adjusted. The incident angles for phase-matching in the bulk must be  $\theta_{1,I} = 88.90$  degrees,  $\theta_{2,I} = 88.98$  degrees, and  $\theta_{3,I} = 88.92$  degrees if the azimuthal angles are zero. The bulk wave then is ‘reflected’ at  $\theta_{4,\mu3} = 89.65$  degrees.

A numerical analysis with three mixed Gaussian beams was performed. As before each Gaussian beam was decomposed into seven plane waves and focused to a 10 micron spot via a 1 meter lens. The center of the mixed beam ‘reflects’ at 89.7 degrees, but has a larger-than-expected angular width. The centers of the incident beams have polar angles identical to those used in the previous example ( $\theta_{1,I} = 88.90$ ,  $\theta_{2,I} = 88.96$ ,  $\theta_{3,I} = 88.91$  degrees). The angular profile of the generated beam shows a bright spot that is off center from the center of the halo. This angular profile is shown in Figure 3.44. The small dark region to the left of center represents the location of the bright spot; predominately

surface-generated light. The large halo shown in the profile is much broader than the surface signal owing to the smaller sensitivity to mismatch. At the location in the beam where the surface intensity peaks, this experiment yields a surface-to-bulk signal ratio of 21, when masking off the halo.

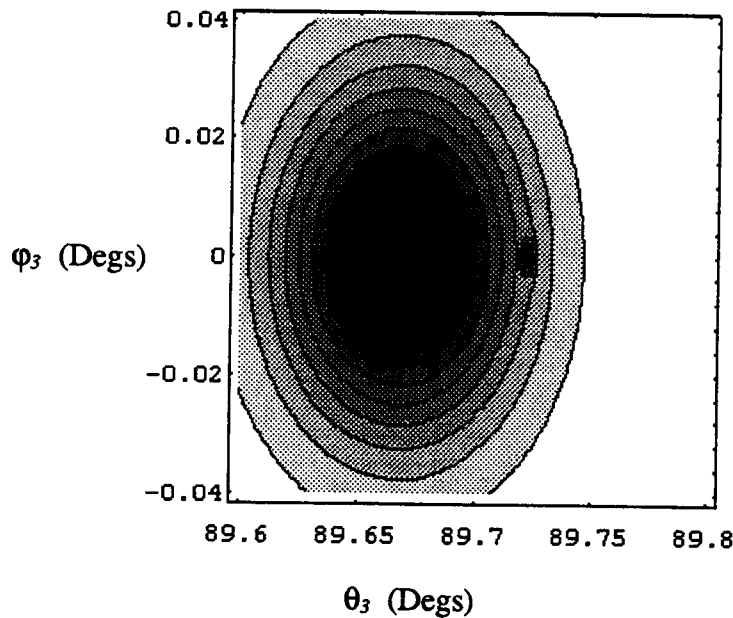


Figure 3.44: A contour plot of the intensity angular profile versus the polar and azimuthal angles. The process is Gaussian beam CARS-like FWM with dispersion.

### 3.8 Conclusions

An exact macroscopic treatment of general optical mixing at a dielectric interface was presented. In this treatment, a model structure (Figure 3.13) was used to obtain general expressions for the bulk and surface contributions to the generated signal. The model structure divides the interfacial region into four distinct regions: a cover, a surface region,

a bulk medium, and the semi-infinite bulk. Using Maxwell's equations, the boundary conditions, and the solution to the inhomogeneous optical wave equation, the generated fields were derived. Using the field expressions for the 'reflected' mixed waves (the RISOM expressions), predictions were made suggesting favourable conditions for RISOM experiments. This was done by assuming a set of physically realistic parameters (indices of refraction and thicknesses of regions, polarization directions, dispersions, etc.) and analyzing the RISOM expressions under various configurations. The next several paragraphs recapitulate these analyses.

The analyses of the RISOM expressions are especially involved. To best examine general mixing processes such as TWM and FWM in media with dispersion, the various pieces to the puzzle of predicting RISOM conditions were isolated and studied individually. As the analysis proceeded, more pieces were added to form the mosaic that describes general reflection geometry OM processes. During the examination of each part of this collage of angular dependencies, layer indices, layer thicknesses, and mismatch; particular attention was paid to conditions which might enhance the surface over the bulk contribution. The numerical analyses of these expressions was first divided into two parts: collinear and non-collinear geometries. Each part was examined for the affects of polarization and degree of dispersion for general and specific OM processes. When favourable conditions were found, actual RISOM experiments are suggested and discussed. Furthermore, when such conditions were predicted, calculations for experiments using Gaussian beams, and not simply plane waves, were performed.

For general collinear optical mixing, including degenerate ( $n$ -th) harmonic generation ( $n$ HG) and TWM or FWM in dispersionless media, were performed first. These calculations laid the groundwork for all subsequent calculations as they exhibit the most basic angular dependencies for reflection geometry OM. For  $S$ -polarizations, reasonably good enrichment of the surface contribution was predicted for near-grazing incidence. The surface-to-bulk signal ratio for an angle of 89 degrees can be expected to be on the

order of 2. Analysis of *S*-wave OM processes in media *with dispersion* showed that its inclusion does not change the angular behaviour of the bulk and surface contributions. The only effect is that the overall intensity is reduced because of the phase-mismatch. It was shown that for *P*-polarization experiments in *dispersionless* media may or may not exhibit a nonlinear Brewster's angle. An investigation of *P*-waves for each of these possibilities was conducted. Scrutiny of *P*-polarization collinear processes in *dispersionless* media where both the surface and bulk have a NLBA suggested that there is no possibility of separating the bulk and surface signals. In this case the NLBA for each wave will be equal. When the polarizations are not induced in the same direction, however, the NLBAs are different. This provides a definitive way to separate the surface or the bulk contributions from the total signal. Similarly, when *dispersion* is included, the NLBAs also differ. Even with small dispersions, the disparity between NLBAs can be experimentally observable. The difference between surface and bulk Brewster's angles is typically 2 to 8 degrees. Utilizing this difference, separation of surface and bulk can be quite favourable; the direct ratio can be 5 to 200, or larger.

Analyses of two non-collinear experiments which intentionally introduce a phase-mismatch were performed. Experimentally, phase-mismatches can be induced by slightly changing the direction of one (or more) of the incident wavevectors from the phase-matched case. The degree of mismatch is left to the discretion of the experimenter. The two experimental designs involve either selecting different polar angles, leaving the azimuthal angles similar, or the converse. When the media are *dispersionless*, the surface and bulk waves are 'reflected' from the structure at the same angle. They may be separated in *intensity* by utilizing a difference of the sensitivity in the phase-mismatch for the two layers. The surface-to-bulk signal ratio can be 2 to 3. With the inclusion of *dispersion*, the bulk and surface contributions may be separated by phase-matching in one layer and not the other. This is possible only if the dispersions for the two layers are different -- the larger the difference the better the separation.

It was shown that in *dispersive media* general non-collinear TWM processes acquire a finite angular separation between the surface and bulk waves. Specific examples of plane wave SFG processes and SFG using Gaussian beams were investigated. The mixed Gaussian beam ‘experiment’ shows an angular profile which has a bright spot surrounded by a much larger dim halo. The bright spot contains mostly surface-generated light. Its width can be 0.2 degrees. The halo is comprised of predominately bulk signal; the halo is on the order of 1.5 degrees in width. The surface-to-bulk ratio can be on the order of 4 for this experiment.

As with TWM, an analyses for general non-collinear FWM in media *with dispersion* was performed. For plane waves, typical angular separations were 0.05 degrees between the surface and bulk. A careful analysis using Gaussian beams was provided. Again, the mixed Gaussian beam shows an angular profile which has a bright spot, the surface signal (typically 0.01 degrees in width), surrounded by a much larger dim halo (the bulk signal) which is on the order of 0.25 degrees in width. The surface-to-bulk ratio can be on the order of 20 for this experiment.

In summary, reflective geometry experiments are straightforward, signals are typically large, and there are no apparent limitations on the type of surface that can be studied. Favourable surface-to-bulk ratios are achieved by making use of differences in the angular and phase-mismatch dependencies of the surface and bulk waves. These dependencies may allow the experimenter to isolate the surface signal in a RISOM geometry.



## 4. CONCLUSIONS

This research was driven by the potential to further develop the tools of optical surface science. The central question of this research has been: can vibrational and electronic surface resonances involving optical transitions be detected on flat surfaces?

Chapter 2 took the approach of manipulating the applied fields in the confines of a planar waveguide to achieve an interference condition. This interference condition is engineered to exclude the generation of signals everywhere except the surface. The research mapped out experimentally relevant parameter space for CARS-like WISOM processes in a waveguide. Non-collinear geometry and phase-matching were considered in the waveguide geometry as a way to increase the overall surface signal and surface-to-bulk signal ratios. In Chapter 3, reflection geometry ISOM was considered. For RISOM experiments, it was found that manipulation of the applied fields can maximize the surface-to-bulk signal ratio. The angular dependence of optical mixing and phase-matching conditions are analyzed as a way to predict enhanced surface-specificity. In order to predict successful RISOM experiments, an entire treatment of optical mixing at a surface that includes both surface and bulk contributions was developed.

### 4.1 Planar Waveguide Geometry Interface-Specific Optical Mixing (WISOM)

As means of observing surface resonances on flat surfaces, Chapter 2 considered optical mixing in waveguide geometry. Using WISOM, the surface region's frequency response, symmetry information, and perhaps (time-wise) transitory behaviour can be investigated. These properties may be used to determine the stoichiometry, adhesion mechanisms, migration times, and reaction properties of species comprising the 'surface'.

A general theoretical treatment for WISOM began with the derivation of the fields in a general planar waveguide. If the index profile of the waveguide is known, the fields can be found by employing Maxwell's equations to find a wave equation. The wave equation for an inhomogeneous, dielectric waveguide was derived in Appendix A. A short discussion illuminates how the fields of a graded-index waveguide may be calculated. The fields for an asymmetric, step index waveguide were found explicitly by considering the Helmholtz equation. The Helmholtz equation was separated into three equations, one for each region of the waveguide. Boundary conditions at each interface were used to derive the field in the three regions. Initially, only plane waves were considered in this solution; finite waves were later considered utilizing superpositions of the derived plane wave solutions. In matching the boundary conditions, a transcendental equation was revealed relating the wavevectors of the various regions. The solution of this transcendental equation exposed the discrete modal nature of the waveguides. The resulting expressions derived for the field amplitude and phase of a guided wave relied only on the assumptions that the indices of refraction are time-independent and that the waveguide is constructed from dielectric materials. The last part of the theoretical discussion expressed the field of the generated plane wave in terms of the parameters of the incident plane waves, the linear indices, and non-linear susceptibilities of the cover, film, and substrate. Interpretative tools that best quantify surface-specificity, the direct ratio (DR) and the cross ratio (CR), were introduced. A Mathematica notebook was written to obtain numerical solutions based on the theoretical development. Numerical solutions for the generated electric field could be found given assumed waveguide index profiles, dimensions, and applied wave frequencies.

Analysis of the relevant parameter space for CARS-like WISOM using infinite plane waves began by assuming a film index and a thickness. The range of waveguide parameters are dictated by what is realizable in experiments. Hence, the range of the index of the film was restricted to 1.60 to 2.10. The thickness range of the film were likewise confined to 600 to 950 nm. The substrate was assumed to be fused silica and its

index was appropriately fixed at 1.47. The cover index also remained fixed throughout the analysis; it was taken as that of the vacuum. The frequencies for the applied waves were chosen with values appropriate for vibrational resonances. The first and third wave frequencies were chosen to be consistent with  $\lambda_1 = \lambda_3 = 560$  nm. The second wave was picked by means of the difference in wavenumbers between the second and first waves. The equivalent difference frequency ranged from 1000 to 3000 wavenumbers (roughly 580 to 670 nm) in steps of  $200\text{ cm}^{-1}$ . All possible mode combinations of these four waves were considered. The values of  $\chi^{(3)}$  for the various regions of substrate, film, and vacuum were taken as unity, except over the cover region 0.5 nm above the film. In keeping with a cover region in resonance with the mixing process,  $\chi^{(3)}$  was taken as 100 for the cover; a rather conservative value considering the typical range is 10 to 100 times the non-resonant value. It should be recalled that no attempt was made to examine possible contributions to an interstitial region between the substrate and film regions. Over the tuning range, the non-linear susceptibilities were considered fixed.

For each waveguide (i.e., choice of thickness and film index) and mode combination, the direct and cross ratios were tabulated; three-dimensional plots were created for ease in interpretation. A series of plots for each mode combination was generated. Each plot in a series represents a unique difference frequency. Although all mode combinations that have the potential to yield favourable WISOM conditions were analyzed, only six were found to be remarkable. These are the six combinations having the largest DRs, CRs, tuning ranges, and best phase-matching options out of the full eighteen possible combinations. These promising mode combinations are: [2,2,0,1], [1,0,1,1], [1,1,1,2], [0,0,0,1], [2,0,2,1], and [1,1,0,1]. Two of these combinations are four-beam geometry combinations. Phase-matching can be achieved with small azimuthal angles when mode combination [1,1,0,1] is employed.

Phase-matching inside a waveguide was found to enhance the overall signal levels of WISOM processes. Moreover, it was discovered that if the dispersions of film and cover

regions differ, the surface-specificity can be very strong. In order to phase-match the non-linear process, the azimuthal angles must be chosen. From the previous set of surface-specificity calculations for the various mode combination and waveguides, it was found that the possibility of phase-matching in the waveguide is quite common. In over 70% of the choices of waveguides, mode combinations, and difference frequencies, phase-matching occurs when the azimuthal angles are approximately perpendicular from one another. Furthermore, in about 18% of the choices, phase-matching occurs when the azimuthal angles were under 3 degrees.

A theoretical development was constructed in order to evaluate phase-matching in the waveguide, especially those of four-beam mode combinations. To best simulate laboratory experiments, three focused Gaussian profile beams coupled into a waveguide were considered. The development began by deriving a mathematical description of a focused Gaussian beam. A numerical analysis of the generated signals for phase-mismatched conditions where the interaction length is large versus the case where phase-matching is achieved was made using the Gaussian beams development. These were accomplished by assuming 1 mm diameter Gaussian beams focused by 1 meter lenses. The direction of the wavevectors for each beam was found for every point in the interaction volume by matrix methods of paraxial optics. Numerical calculations were done for many modes, difference frequencies, and waveguides. Two situations were identified: the case where the azimuthal angles for zero mismatch are all small and the case where azimuthal angles are essentially perpendicular. The variance of intensity versus phase-mismatch was scrutinized for both situations. When the azimuthal angles for zero mismatch are all small, there is a large interaction volume for phase-matched conditions. It was shown that when the applied beams are collinear, the phase-mismatch is large and the generated intensity is low. In the example given, the ratio of intensities between the phase-matched angles versus collinear beams was 2.97. When the phase-matching azimuthal angles are small, it is advantageous to phase-match the OM process rather than to maximize the interaction volume by using collinear beams. Also seen in the example is

the *sinc*-like dependence of the phase-mismatch on intensity. This dependence is modulated by a high-frequency  $\cos^2$ -like behaviour on the phase-mismatch. From the analysis it was found that it is more common for WISOM processes to be phase-matched when the beams are nearly perpendicular. This is the second situation: the interaction volume for phase-matching is small compared to the collinear application of the input beams. Despite the shorter interaction length, the phase-matched case generally provides signals two to three times larger. For the example cited, the intensity ratio of the phase-matched, small interaction region condition versus the collinear case was 2.8. Moreover, when the dispersions of the two layers are different, even for a dispersion of one part in  $10^3$  for the surface and twice that for the film, there is an additional enhancement in the surface-to-film ratio of  $10^3$  to  $10^4$ .

It was concluded that there are benefits to employing phase-matching and utilizing the four-beam mode combinations. To reap these benefits, however, four couplers are required to couple the beams to the waveguide. Traditionally, researchers have used prism couplers which are very cumbersome. Furthermore, it is difficult to secure two functional couplers to a waveguide; securing four couplers seems impractical. Consequently, efforts were made to improve coupler design and construction for asymmetric planar waveguides.

Achromatic waveguide couplers were considered as an alternative. These are modified prism couplers such as those proposed by Spaulding and Morris. Their use simplifies coupling by making the coupling efficiency less sensitive to the frequency and coupling angle of the beam. Unaddressed is the problem of the variability of coupling efficiency inherent between one prism coupler and another.

Integrated grating couplers were investigated as possible solutions to the coupling problem. For purposes of general CARS-WISOM work, four grating couplers per waveguide must be fabricated. Simple calculations showed that the ideal grating spacing

is between 450 and 500 nm. Experimental attempts with constructing reproducible gratings on substrates were done. Fabrication techniques for line gratings were outlined. The procedure involved spin-coating a UV-sensitive photoresist on a substrate. Exposure of a cosine-squared intensity pattern was applied to the photoresist. Grating spacings (in the photoresist) were produced ranging from 477 to 2785 nm. These gratings were found to have very reproducible spacings and diffraction efficiencies. Furthermore, when producing gratings in the photoresist, the success rate was very good. However, only one in twenty gratings were successfully etched into the substrate.

One proposed solution to the coupling problem involved inducing gratings in a photorefractive medium embedded in the waveguide. An active sinusoidal index grating would be induced by applying two interfering probe beams. The WISOM beams would then be coupled into the waveguide using these active gratings. The efficiency of these gratings is typically not exceptional. The sacrifice of efficiency, however, for the attractive ability to alter the orientation and spacing of the gratings may be acceptable.

It was suggested that ring gratings be constructed in place of linear gratings to couple to the waveguides. Since the phase-matching angles are very different for different mode combinations and difference frequencies, this configuration seems natural. Although no gratings of this type have been demonstrated, construction methods were proposed. Fabricating these structures could be done using the same methods for making linear gratings. However, the interference pattern used to expose the ring grating must be created in different ways. Two methods for exposing the ring pattern were outlined. The first suggested construction method for ring gratings involved masked diffraction from a circular aperture. Using a 356 nm source, a circular aperture 1.9 mm in diameter placed 5 mm from the photoresist would produce a 468 nm ring grating. The center 3 mm of the diffraction pattern would be masked off. The spacing would vary only 2% over 5000 rulings for this grating. A longer exposure is required since most of the light is in the first few orders of the diffraction pattern which are blocked to create a clear aperture. The

exposure time for the ring grating was calculated to be over 13 hours. Vibrations would likely destroy the image in the photoresist as it was being exposed. Another alternative to create the ring pattern directly utilizing a diffractive optic was conceived. It was shown that a diffractive optic for a ring grating can be obtained by taking the inverse Hankel transform of the desired pattern. A diffractive optic was numerically designed that would produce a circular pattern with 5000 concentric ring fringes spaced by 450 nm with the first fringe having a radius of 1.5 mm.

These proposed solutions to the problem of radiative coupling to planar waveguides are incomplete, as no working structures have been produced. Continued research and development into the fabrication of alternative grating couplers, such as ring gratings and integrated non-linear grating couplers, is needed.

The WISOM geometry is fraught with experimental difficulties and provides limited surface-specificity. A major limitation of WISOM is that only surfaces of thin dielectric films can be studied. Scanning through the resonances of a surface using the WISOM technique may require a series of waveguides due to the frequency sensitivity of the interference condition. However, ways of increasing the signal levels and enhancing the interference conditions were investigated. The experiments that produce these enhancements were discovered to be experimentally involved. These limitations led to the investigation of another ISOM geometry discussed in Chapter 3.

## **4.2 Reflection Geometry Interface-Specific Optical Mixing (RISOM)**

Chapter 3 took a very different approach, suggesting that reflection geometry could be employed in an optical mixing process to provide good surface-to-bulk signal ratios. The advantages of a reflective geometry experiment is that it is straightforward and phase-matching is easy to achieve. Signals are typically large and there are no apparent

limitations on the type of surface that can be studied. The origin of favourable signal-to-noise ratios in RISOM experiments is not directly obvious.

In the RISOM research disseminated here, a macroscopic model of OM at an interface was used to obtain the non-linear Fresnel equations. Why a microscopic model was not used is explained in the difficulty of calculation; microscopic model calculations are rather involved. Since the parameters of RISOM were wholly unknown and many calculations were expected, an abbreviated approach was sought. Hence, a well-established macroscopic model was chosen and modified to predict surface-to-bulk signal ratios. In developing this exact macroscopic model, careful attention was paid to phase-matching in both the surface and bulk regions. This allows for the characterization of experimental RISOM parameters when dispersion and intentionally introduced phase-mismatches are to be considered. The next several paragraphs detail the Bloembergen model, the macroscopic model developed to predict surface-to-bulk signal ratios, and the justifications to use a macroscopic instead of a microscopic one.

Ju's microscopic calculation of the optical response lends justification to the use of a macroscopic model in place of a microscopic one. He calculated the optical response of a semi-infinite medium consisting of discrete point-dipoles. Certain surface effects are evident in Ju's analysis. However, he showed that when these surface effects are averaged, the macroscopic result is reproduced. In reviewing calculations based on microscopic models, it became clear that embarking on the microscopic calculations for RISOM conditions without prior direction would be an inefficient use of resources. Using a macroscopic model, most of the calculations to map out a RISOM parameter space can be done analytically; predictions are much simpler to formulate.

An exact macroscopic treatment of general optical mixing at a dielectric interface based on Bloembergen's example was developed. This model breaks the interfacial region into *four* distinct regions: the vacuum, a surface region, a bulk median, and the semi-



infinite bulk. The additional layer is a modification of Bloembergen's model; it was added between the surface slab and the bulk. This layer has the non-linear properties of the bulk which are, in general, different from the surface. The linear index of refraction of the bulk median and the semi-infinite bulk are dissimilar. An attempt was made to include phase-mismatching by integrating the generated fields over the interaction region. From these results, RISOM experiments were suggested in which the surface region may be phase-matched, but phase-mismatched in the bulk. The solutions for the fields generated in each region presented in Chapter 3 were quite general: they were found without exact knowledge of the induced non-linear polarization. No assumptions concerning the particular optical mixing process were necessary. Hence, the reflection coefficients for surface and bulk were derived for *general optical mixing*. From the equations for general mixing, a simple substitution can be made to predict RISOM for specific processes such as CARS-like FWM, SHG, and TWM.

There are two clear predictions made by the macroscopic model. First, an enhancement of the surface over the bulk signal due to phase-mismatch is expected if dispersion of the surface and bulk (median) are different. The second prediction involves the separation of the surface and bulk signals due to dissimilar NLBAs (for *P*-polarization experiments). Any difference in dispersion leads to the condition where the NLBA for the bulk is unequal to the NLBA for the surface.

The numerical analyses of the RISOM expressions examine mixing processes such as TWM and FWM in media with dispersion. To understand the complexities of the RISOM solutions, increasingly specific situations were individually studied. Particular attention was paid to conditions which might enhance the surface over the bulk contribution. The contributions of the surface and bulk regions to the electric fields were numerically calculated for each of case of polarization, degree of dispersion, and geometry of the waves (collinear or non-collinear). The analyses were first divided into two parts: collinear and non-collinear geometries. Each geometry was examined under different

conditions of polarization and degree of dispersion for general and specific OM processes. When favourable conditions were found, RISOM experiments were suggested. Furthermore, when such conditions were predicted, calculations for experiments using focused Gaussian beams in place of plane waves were performed. These beams were mathematically represented by a decomposition of seven plane waves focused to a 10 micron spot using a 1 meter lens. The next several paragraphs review these analyses.

Generally throughout the analysis, calculations were done with a consistent set of fixed parameters. The index of refraction of the cover region is the index of the vacuum. The index of the surface, bulk, and bulk median are seen as scaling parameters. They are set to: 1.50 for the surface index, 1.55 for the bulk median index, 1.56 for the bulk index. The layer thicknesses of the surface and bulk median are both set to 5 nm, much smaller than the wavelength of light.

A collinear *S*-wave RISOM experiment was predicted for near-grazing angles of incidence. For instance, the surface-to-bulk signal ratio for an angle of 89 degrees from non-dispersive media can be expected to be on the order of 2. Analysis of *S*-wave *n*HG processes in media *with dispersion* showed that the angular behaviour of the bulk and surface contributions remained unchanged.

Examination of the RISOM expressions yielded that collinear optical mixing performed in *dispersionless* media for *P*-polarizations may or may not exhibit a non-linear Brewster's angle (NLBA). When a *P*-polarization collinear *n*HG process in *dispersionless* media has a NLBA in both media, analysis of collinear SHG suggested that there is no possibility of separating the bulk and surface signals. With small dispersions, however, the surface and bulk NLBAs differ. Utilizing this difference, the surface-to-bulk signal ratio can be from 5 to 200.

Analyses of two experiments which intentionally introduced a phase-mismatch were performed. Experimentally, phase-mismatches can be produced by using a non-collinear geometry. Using the non-collinear geometry, the experimenter has control of the degree of phase-mismatch. The two experimental designs involve either selecting different polar angles, leaving the azimuthal angles similar, or the converse. Two main conclusions come of these analyses. First, when the media are *dispersionless*, the surface and bulk waves are ‘reflected’ from the structure at the same angle. They can be separated in *intensity* by phase-mismatching the mixing process. The other conclusion is that, with the inclusion of *dispersion*, the bulk and surface waves are separated in *angle*.

The first non-collinear reflection OM experiment introduces a mismatch by adjusting only the polar angles of the incident waves. Analysis was simplified by considering only SHG. The NLRCs for *S*-wave and *P*-wave (without a NLBA) non-collinear SHG were found to be similar. The angular dependence was found to be complicated and was different for the surface and bulk contributions. With azimuthal angles of the fundamental waves equal, the polar angle dependence on the sensitivity of bulk and surface amplitudes due to the *difference* in polar angles was investigated. The polar angle mismatch dependence (or ‘mismatch sensitivity’) was mapped out. It was discovered that the *surface* mismatch sensitivity increases simply with polar angle. However, the mismatch for the *bulk* is very sensitive at a certain polar angle. An approximate expression for this polar angle was derived. At this angle, when the difference between polar angles is zero, the signal is predominately from the bulk. In contrast, when difference between polar angles is a few milli-degrees, the bulk is highly attenuated; typical surface-to-bulk ratios (using plane waves) are on the order of 2. Estimates for a Gaussian beam SHG experiment shows that the SHG beam has a profile enriched in surface contribution due to the different surface and bulk phase-mismatch sensitivities. The intensity at the center of the SHG beam is a mixture of surface- and bulk-generated light. The angular width of this center portion may be 0.02 degrees (full-width half-maximum). The outer portion of the SHG beam has a much lower intensity, but is due primarily to surface-generated light.

This RISOM experiment may yield a surface-to-bulk signal ratio of 10. If the NLBA happens to be near the angle at which the bulk is most sensitive to mismatches, then the NLBA can be used to further enhance the bulk signal over the surface contribution by an additional factor of 10. Unfortunately, the total signal strength suffers. Finally, it was argued that the mismatch is further increased when *dispersion* is included. When dispersion is different, the separation of bulk and surface signals by phase-mismatch is more pronounced.

The second 'experiment' introduces a mismatch by adjusting only the azimuthal angles of the incident waves. Again, analysis was simplified by considering only SHG. The NLRCs for non-collinear SHG were found to be similar for both *S*- and *P*-waves (without a NLBA). In *dispersionless* media, it was found that the bulk wave is rather insensitive to the phase-mismatch as compared to the dependence of the surface wave. Unfortunately, this leaves no way to attenuate the bulk signal over the surface wave. The reverse is possible, however, providing a way to observe the isolated bulk signal. From observations like this, bulk non-linear susceptibility tensor elements may be examined.

It was argued that general TWM in media with *dispersion* was similar to SHG, with the difference that the process is phase-matched when the incident angles are *not* collinear. As with SHG, altering the directions of the incident waves from their minimally phase-mismatched positions introduces phase-mismatch that can be used to separate the surface and bulk contributions. For *P*-polarization experiments, if a NLBA exists then further enhancements in the signal ratio are possible. Specific examples of plane wave SFG processes and SFG using Gaussian beams were investigated using incident frequencies of 500 and 600 nm. The SFG Gaussian beam 'experiment' showed that the generated beam has an angular profile which has a bright spot surrounded by a much larger dim halo. The halo, with a width on the order of 0.5 degrees, is comprised of predominately bulk signal. The bright spot contains mostly surface-generated light. Its width is typically 0.01 degrees. The surface-to-bulk ratio can be on the order of 5 for this experiment. By

increasing the mismatch sensitivity of the bulk, the halo can be broadened; this can improve the signal-to-bulk ratio to  $10^8$ .

An analysis for general non-collinear FWM in media *without dispersion* was performed. Three incident waves of wavelengths 500, 550, and 600 nm were used. It was found that the addition of the third incident wave made introducing phase-mismatch sensitivity of the bulk without effecting the surface sensitivity more likely. Non-collinear geometry FWM in media with *dispersion* was analyzed. For plane waves, typical angular separations were 0.05 degrees between the surface and bulk. An analysis using focused Gaussian beams showed the generated beam has an angular profile with a bright surface-dominated spot (typically 0.01 degrees in width), in the field of a larger dim halo (the bulk signal) which is on the order of 0.25 degrees in width. The surface-to-bulk ratio can be on the order of 20 for this experiment. Increasing the mismatch sensitivity of the bulk can improve the signal-to-bulk ratio and can be used to obtain a surface-to-bulk ratio of  $10^{11}$ .

In summary, simple reflection geometry OM experiments can be employed that are able to distinguish a surface component to a signal that is composed mostly of bulk contributions. Reflective geometry experiments are straightforward, signals are typically large, and there are no apparent limitations on the type of surface that can be studied. This is to be contrasted with the waveguide geometry, where the incident and generated waves must be coupled into a guiding structure. Utilizing the differences in the angular phase-mismatch dependencies of the surface and bulk waves, favourable surface-to-bulk ratios can be achieved. These dependencies were investigated in detail. It was found that reasonable surface-to-bulk signals are possible and several experimental designs are suggested. Future microscopic calculations are suggested in circumstances where the macroscopic model is in question. Also, the parameter space was significantly narrowed allowing future researchers employing microscopic models to investigate specific areas of interest.

## BIBLIOGRAPHY

- [1] W. M. Hetherington III, E. W. Koenig, W. M. K. P. Wijekoon, *Chem. Phys. Lett.* 134, 3 (1987).
- [2] W. M. K. P. Wijekoon, Z. Z. Ho, W. M. Hetherington III, *Chem. Phys. Lett.* 86, 8 (1987).
- [3] W. M. Hetherington III, Z. Z. Ho, E. W. Koenig, G. I. Stegeman, R. M. Fortenberry, *Chem. Phys. Lett.* 128, 2 (1986).
- [4] P. Franken, A. E. Hill, C. W. Peters, G. Weinreich, *Phys. Rev. Lett.* 7, 118 (1961).
- [5] M. Bass, P. Franken, A. E. Hill, C. W. Peters, G. Weinreich, *Phys. Rev. Lett.* 8, 18 (1962).
- [6] J. A. Giordmaine, *Phys. Rev. Lett.* 8, 19, (1962).
- [7] P. D. Maker, R. W. Terhune, M. Niseoff, C. M. Savage *Phys. Rev. Lett.* 8, 21 (1962).
- [8] R. W. Terhune, P. D. Maker, M. Niseoff, C. M. Savage *Phys. Rev. Lett.* 8, 404 (1962).
- [9] R. M. Corn, D. A. Higgins, *Chem. Rev.* 94, 107 (1994).
- [10] N. Bloembergen, Non-linear Optics, Reprinted Edition, Addison-Wesley, 1992. See Chapter 4.
- [11] Amnon Yariv, Quantum Electrodynamics, John Wiley and Sons, Inc., Second edition, 1975.
- [12] Y. R. Shen, Optical Second Harmonic Generation at Interfaces, *Ann. Rev. Phys. Chem.* 40, 327-350 (1989).
- [13] N. Bloembergen, R. K. Chang, C. H. Lee. Second-harmonic Generation of Light in Reflection from Media with Inversion Symmetry, *Phys. Rev. Lett.* 16, 986-989 (1966).
- [14] C. H. Lee, R. K. Chang, N. Bloembergen. Nonlinear Electreflectance in Silicon and Silver, *Phys. Rev. Lett.* 18, 167-170 (1967).
- [15] N. Bloembergen, R. K. Chang, S. Jha, C. H. Lee, Optical Second Harmonic Generation in Reflection from Materials with Inversion Symmetry, *Phys. Rev.* 174, 813-822 (1968).
- [16] F. Brown and M. Matsuoka. Effect of Adsorbed Surface Layers on Second-harmonic Light from Silver, *Phys. Rev.* 185, 985-987 (1969).

- [17] J. M. Chen, J. H. Bower, C. S. Wang, C. H. Lee, Optical Second Harmonic Generation from Submonolayer Na-covered Ge Surfaces, *Opt. Comm.*, 9, 132-134 (1973).
- [18] Li-Qiong Wang, Donald R. Baer, Mark H. Engelhard, Ashley N. Shultz, The Adsorption of Liquid and Vapor Water on  $\text{TiO}_2$  (110) Surfaces: The role of defects, *Surface Science*, 344:237-250 (1995).
- [19] Li-Qiong Wang, Ashley N. Shultz, Donald R. Baer, Mark H. Engelhard, Interactions of Small Molecules with  $\text{TiO}_2$  (110) Surfaces: The Role of Defects, *J. Vac. Soc. Technol A* 14, 3, 1532 (1996).
- [20] Ashley N. Shultz, Winyann Jang, W.M. Hetherington III, D. R. Baer, Li-Qiong Wang, M.H. Engelhard, Comparative Second Harmonic Generation and X-ray Photoelectron Spectroscopy Studies of the UV Creation and  $\text{O}_2$  Healing of  $\text{Ti}^{3+}$  Defects on (110) Rutile  $\text{TiO}_2$  Surfaces, *Surface Science* 339, 114-124 (1995).
- [21] Li-Qiong Wang, K.F. Ferris, A.N. Shultz, D.R. Baer, M.H. Engelhard, Interactions of  $\text{HCOOH}$  with Stoichiometric and Defective  $\text{TiO}_2$  (110) Surfaces, to be published.
- [22] Chang-Yuan Ju, Theory and Application of Optical Second Harmonic Generation on Dielectric Surfaces, PhD Thesis, Oregon State University, 1994.
- [23] N. Bloembergen, P. S. Pershan, *Phys. Rev.* 128, 606 (1962).
- [24] J. Ducuing, N. Bloembergen, *Phys. Rev. Lett.* 10, 474 (1963).
- [25] N. Bloembergen, Chang, R. K. Jha, S. S. C. H. Lee, *Phys. Rev.* 174, 813 (1968).
- [26] W. M. Hetherington III, N. E. van Wyck, G. I. Stegeman, R. M. Fortenberry, *Opt. Lett.*, 9, 88 (1984).
- [27] W. M. Hetherington III, N. E. van Wyck, E. W. Koenig, G. I. Stegeman, R. M. Fortenberry, *Proc. SPIE* 620 (1986).
- [28] It might be interesting to choose materials that exhibit the D.C. Pockels or D.C. Kerr effect. This has been suggested by W. M. Hetherington during private communications.
- [29] Private communications (in 1994) with Dr. Tom Plant, Professor of Electrical and Computer Engineering, Oregon State University.
- [30] Z. Z. Ho, W. M. K. P. Wijekoon, E. W. Keonig, W. M. Hetherington, III, *J. Phys. Chem.*, 86, 8, 4384 (1987).



- [31] Z. Z. Ho, W. M. K. P. Wijekoon, E. W. Keonig, W. M. Hetherington, III, *J. Phys. Chem.*, 91, 757 (1987).
- [32] R. Fedosejevs, M.J. Brett, *Appl. Opt.*, 28, 10, 1877 (1989).
- [33] X. Mai, R. Moshrefzadeh, U. Gibson, G. Stegeman, C. Seaton, *Appl. Opt.*, 24, 19, 3155 (1985).
- [34] Y. Nakano, K. Tada, *Opt. Lett.*, 13, 1, 7 (1988).
- [35] D. Marcuse, Light Transmission Optics, New York, Van Nostrand Reinhold 1972.
- [36] L. M. Brekchovskich, Waves in Layered Media, New York, Academic 1960.
- [37] J. B. Keller, J. S. Papadakis (Ed.), Wave Propagation and Underwater Acoustics, Springer Lecture Notes in Physics, vol.70. Berlin, Springer Verlag 1977.
- [38] A. V. Gurevich, E. E. Tsedilina, Superfar Propagation of Short Radiowaves, Moscow, Nauka 1979.
- [39] J. A. Arnaud, Beam and Fiber Optics, New York, Academ. Press 1976.
- [40] M. S. Sodha, A. K. Ghatak, Inhomogeneous Optical Waveguides, New York, Pledum Press 1977.
- [41] V. M. Babitch, V. S. Buldyrev, Asymptotic Methods in Problems of the Diffraction of Short Waves, Moscow, Nauka 1972.
- [42] V. A. Borovikov, B. E. Kinber, Geometrical Theory of Diffraction, Moscow, Sviaz 1978.
- [43] Kravisov, A. Yu, Orlov, L. Yu, Geometrical Optics of Inhomogeneous Media, Moscow, Nauka 1980.
- [44] V. P. Maslov, Operator Methods, Moscow, Nauka 1973.
- [45] V. P. Maslov, M. V. Fedorink, Quasiclassical Approximations for Equations of Quantum Mechanics, Moscow, Nauka 1976.
- [46] B. R. Vainsberg, Maslov's Canonical Operator: Theory of Wave Propagation in Inhomogeneous and Nonlinear Media, Moscow, Institute of Radiotechnics and Electronics 1979.
- [47] L. B. Felsen, Evanescent Waves. *J. Opt. Soc. Amer.* 66, 751 -760 (1976).

- [48] Kravtsov, A. Yu, Complex rays and Complex Caustics, *Izv. Vuz. Radiofiz.* (USSR) 10, 1283-1304 (1967) .
- [49] W. Y. D. Wang, G. A. Deschamps, Application of Complex Ray Tracing to Scattering Problems. *Proc. IEEE* 62, 1541-1551 (1974).
- [50] V. M. Babitch, I. A. Molotkov, M. M. Popov, Gaussian Beams Localized in the Vicinity of the Lines of the Solutions and Their Applications, Preprint No. 24., Moscow, Institute of Radiotechnics and Electronics 1984.
- [51] V. M. Babitch, T. F. Pankratova, Problems of Mathematical Physics, Leningrad, LGU 1973, pp. 2-20.
- [52] B. Z. Katzenelenbaum, Theory of Nonuniform Waveguides with Slowly Varying Parameters, Moscow, Acad. Sciences of the USSR 1961.
- [53] V. V. Shevchenko, Continuous Transmissions in Optical Waveguides, Boulder, Colorado, Golem Press 1977.
- [54] V. A. Borovikov, Higher Types of Waves in Slowly Irregular Waveguides. *Radiotekhnika i Elektronika* (USSR) 23, 1365-1376 (1978).
- [55] V. A. Borovikov, A. V. Popov, Propagation of Waves in Slowly Nonuniform Multimode Waveguides. Direct and Inverse Problems of Diffraction. Moscow, Institute of Radiotechnics and Electronics Acad. Sci. USSR 1979, pp. 157-266.
- [56] A. V. Popov, Short Wave Asymptotic of a Normal Wave in a Nonuniform Waveguide, *Radiotekhnika i Elektronika* (USSR) 22, 1577-1582 (1977).
- [57] M. D. Feit, J. A. Fleck, Jr., Light Propagation in Graded-index Optical Fibers. *Appl. Opt.* 17, 3990-3998 (1978).
- [58] M. A. Leontovich, V. A. Fock, Solution of the Problem of Propagation of Electromagnetic Waves Along the Surface of the Earth by Using the Method of the Parabolic Equation. *Zh. Eksper. Teor. Fiz.* (USSR) 16, 557-573 (1946).
- [59] E. A. Polyanskii, Relationship Between Solutions of the Helmholtz and Schrodinger Equations, *Sov. Phys. Acoust.* (USA) 20, 90 (1974).
- [60] E. A. Polyanskii, Method for the Correction of Solutions of the Parabolic Equation for an Inhomogeneous Waveguide, Moscow, Nauka 1985.
- [61] V. A. Fock, Problems of Diffraction and Propagation of Electromagnetic Waves, Moscow, Sov. Radio 1970.

- [62] S. M. Rytov, Kravtsov, A. Yu, Tatarskii, V. I., Introduction to Statistical Radiophysics, Moscow, Nauka 1978.
- [63] V. I. Tatarskii, Propagation of Waves in a Turbulent Atmosphere, Moscow, Nauka 1967.
- [64] Sergej G. Krivoslykov, Quantum-Theoretical Formalism for Inhomogeneous Graded-Index Waveguides, Akad. Verl., Berlin, 1994.
- [65] G. I. Stegeman, R. M. Fortenberry, R. Moshrefzadeh, W. M. Hetherington III, N. E. van Wyck, J. E. Sipe, *Opt. Lett.*, 8, 295 (1983).
- [66] M. T. Wlodarczyk, S. R. Seshadri, Analysis of Grating Couplers for Planar Waveguides, *J. Appl. Phys.*, 58, 1 (1985).
- [67] J. W. Goodman, Introduction to Fourier Optics, McGraw Hill, 1968.
- [68] K. E. Spaulding, G. M. Morris, Achromatic Waveguide Couplers, *J. of Lighthwave Tech.*, 10, 10 (1992).
- [69] M. Xu, R. Moshrefzadeh, U. J. Gibson, G. I. Stegeman, C. T. Seaton, Simple Versatile Method for Fabricating Guided-Wave Gratings, *Appl. Opt.*, 24, 19, 3155-3161 (1985).
- [70] R. Moshrefzadeh, B. Svensson, M. Xu, C. T. Seaton, G. I. Stegeman, Chirped Grating for Efficient Coupling into Nonlinear Waveguides, *Appl. Phys.*, 51, 6, 390-391 (1987).
- [71] R. Fedosejevs, M. J. Brett, Direct Formation of Grating Structures on Silicon using KrF Laser Radiation, , *Appl. Opt.*, 24, 10, 1877-1880 (1989).
- [72] Y. Nakano, K. Tada, In situ Monitoring Technique for Fabrication of High-quality Diffraction Gratings, *Opt. Lett.*, 13, 1, 7-9 (1988).
- [73] Amnon Yariv, Quantum Electrodynamics, John Wiley and Sons, Inc., Second Edition, 1975, pages 522-523.
- [74] J. D. Jackson, Classical Electrodynamics, Second Edition, John Wiley & Sons, 1975.
- [75] J. E. Sipe, *J. Opt. Soc. Am. B* 4, 481 (1987).
- [76] V. Mizrahi, J. E. Sipe, *J. Opt. Soc. Am. B* 5, 660 (1988).
- [77] J. A. Armstrong, N. Bloembergen, J. Ducuing, P. S. Pershan, *Phys. Rev.* 127, 1918 (1962).

- [78] Y. R. Shen, The Principles of Non-Linear Optics, John Wiley & Sons, 1984.
- [79] P. A. M. Dirac, *Proc. Roy. Soc. (London)* A114, 143, 710 (1927).
- [80] M. Goeppert-Mayer, *Ann. Physik* 9, 273 (1931).
- [81] J. Frenkel, Wave Mechanics, Advanced General Theory.
- [82] Th. Neugebauer, *Acta Phys. Acad. Sci. Hung.* 10, 221 (1959).
- [83] D. A. Kleinman, *Phys. Rev.* 125, 87 (1962).
- [84] R. Braunstein, *Phys. Rev.* 125, 475 (1962).
- [85] W. C. Hennenberger, *Bull. Am. Phys. Soc.* 7, 14 (1962).
- [86] H. A. Kramers, Quantum Mechanics, North-Holland Publishing Company, Amsterdam, 1957, pp. 482-489.
- [87] J A Armstrong, N. Bloembergen, J. Ducuing, P. S. Pershan, *Phys. Rev.*, 127, No. 6, 1918 (1962).
- [88] N. Bloembergen and P. S. Pershan, *Phys. Rev.* (1963).
- [89] H. A Lorentz, The Theory of Electrons, B. G. Teubner, Leipzig, 1909.
- [90] L. Rosenfeld, The Theory of Electrons , North-Holland Publishing Company, Amsterdam, 1951.
- [91] O. Klein, *Z. Physik* 40, 407 (1927).
- [92] S. Weinbern, The Quantum Theory of Fields, Cambridge University Press, 1995, pp. 142-143.
- [93] D. P. Craig, T. Thirunamachandran, Molecular Quantum Electrodynamics, Academic Press, 1984, pp. 84-89.
- [94] J. D. Jackson, Classical Electrodynamics, John Wiley & Sons, Second Edition, 1975, pp. 412.
- [95] R. S. Burkey, C. D. Cantrell, *J. Opt. Soc. Am. B* 1, No. 2, 169 (1984).

## APPENDICES

## APPENDIX A

This appendix is devoted to determining the solutions to the generated fields due to a nonlinear source term in an asymmetric waveguide structure as described in Chapter 2. This theoretical development is aimed at deriving ratios that can be used to quantify the degree of surface-specificity in a WISOM experiment. The general development in this appendix expresses the field of the observed wave in terms of the parameters of the incident plane waves (frequency and direction), the linear indices and nonlinear susceptibilities of the cover, film, and substrate regions. Only transverse electric (TE) waveguide modes are considered here.

This appendix has several sections. First, the wave equation is derived for a general waveguide of unknown index gradient. Although the possibility of a graded-index waveguide is not discussed in this work explicitly, the results of this section should prove useful for other researches. The following section is devoted obtaining solutions to the wave equation in the case of an asymmetric step index waveguide. A third section deals with the intensity, phase-mismatch, and quantitative tools used in the numerical analysis of WISOM.

### **A.1 The Wave Equation and Boundary Conditions for Inhomogeneous Waveguides**

There is an unfortunate confusion brought on by the need for multiple sets of coordinate systems describing the directions of the wavevectors and fields. In the end, the choice to define these systems makes the analysis of focused Gaussian beams in a waveguide less difficult. Each beam inside a waveguide is given by its own coordinate system. This makes the describing focused Gaussian beams (in terms of superpositions of

plane waves) more straightforward. The z-axes are taken to be along the central propagation direction of each beam. Calculations involving the culmination of all beams requires a universal waveguide coordinate system. For instance, in determining the nonlinear induced polarization, the waveguide coordinate system is used. The waveguide coordinate system is shown in Figure 2.4. The first beam, the beam that carries  $\omega_1$ , defines the z-axis of the waveguide system. The origin is placed in the center of the film region of the waveguide. Note that each beam's coordinate system and the waveguide system share a common origin. When phase-matching is considered, this common origin will become the crossing center of the beams, that is, the center of the interaction volume. Coordinate transformations between the waveguide (unprimed) system and each of the (primed) beam systems must be obtained. To describe the directions of the waves *outside* the waveguide, before they are coupled or after they are decoupled from the waveguide, a new set of direction angles is defined. Outside the waveguide each wave  $i$  ( $i = 1, \dots, 4$ ) is described by its field  $E_i$  and wavevector  $k_i$ . The direction of  $k_i$  are defined by  $k_i = [|k_i|, \vartheta_i, \alpha_i]$ , where the angle  $\vartheta_i$  is the angle from the x-axis of the waveguide's coordinate system to the wavevector (analogous to the polar angle of a spherical coordinate system). The angle  $\alpha_i$  from the z-axis to the wavevector is analogous to an azimuthal angle. This arrangement is pictured in Figure 2.10.

The frequencies and angular parameters of the waves are referred to using the following convention: all the frequencies for mode combination  $[v_1, v_2, v_3, v_4]$  are represented by  $[\omega_1, \omega_2, \omega_3, \omega_4]$ , the set of 'polar' angles by  $[\vartheta_1, \vartheta_2, \vartheta_3, \vartheta_4]$ , and the 'azimuthal' direction angles by  $[\alpha_1 \equiv 0, \alpha_2, \alpha_3, \alpha_4]$ .

To find the contributions to the mixing process from the cover, film, and substrate, the induced polarizations in each region must be found. This requires that each of the applied fields be known. The fields are found by considering Maxwell's equations and applying

the boundary conditions at each interface. Here the  $i$ -th beam coordinate system is used. Only plane waves are considered in this solution. Recall that in the absence of sources, Maxwell's equations are

$$\begin{aligned}\bar{\nabla} \times \bar{E} &= -\frac{1}{c} \partial_t \bar{B}, & \bar{\nabla} \times \bar{H} &= \frac{4\pi}{c} \bar{J} + \frac{1}{c} \partial_t \bar{D}, \\ \bar{\nabla} \cdot \bar{D} &= 4\pi\rho, & \bar{\nabla} \cdot \bar{B} &= 0.\end{aligned}$$

For a dielectric medium, the true charge density,  $\rho$ , is zero everywhere; the true charge density does not include the polarization charge density. Likewise, the current density,  $J$  is zero. Thus, Maxwell's equations in a dielectric medium using Gaussian units are:

$$\begin{aligned}\bar{\nabla} \times \bar{E} &= -\frac{1}{c} \partial_t \bar{B}, & \bar{\nabla} \times \bar{H} &= \frac{1}{c} \partial_t \bar{D}, \\ \bar{\nabla} \cdot \bar{D} &= 0, & \bar{\nabla} \cdot \bar{B} &= 0.\end{aligned}$$

In general, all the fields are complex and are functions of space ( $x, y, z$ ) and time. Irregardless of the nature of the fields, the general boundary conditions at an interface can be derived. Recall that the coordinate system has its  $z$ -direction parallel to the direction of propagation in the waveguide. Consider a Gaussian volume across an interface as shown in Figure A.1. The integral of Coulomb's law over that volume is

$$\int_v \bar{\nabla} \cdot \bar{D} \, dv = \int_v 0 \, dv = 0.$$



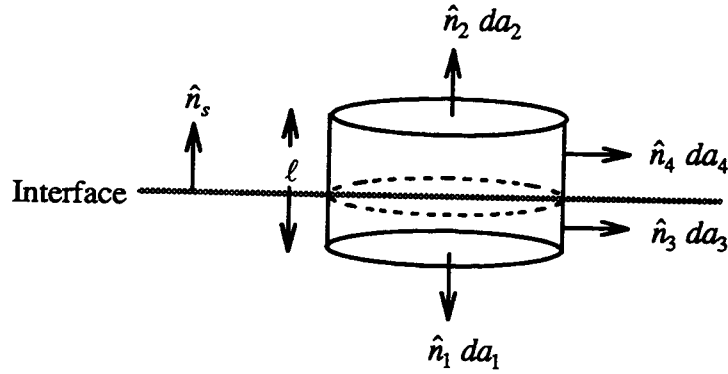


Figure A.1: A cylindrical Gaussian volume straddling an interface.

Applying the divergence theorem, where  $\hat{n}_s$  is the outward (i.e., upward) going surface normal, the integral becomes

$$\int_V \bar{\nabla} \cdot \bar{D} \, dv = \int_S \bar{D} \cdot \hat{n}_s \, da = 0.$$

With  $\bar{D}_1$  the displacement field *below* the interface and  $\bar{D}_2$  displacement field *above* the interface, the integral over the surface  $da = da_1 + da_2 + da_3 + da_4$  becomes

$$\int_V \bar{\nabla} \cdot \bar{D} \, dv = \int_{s_1} \bar{D}_1 \cdot \hat{n}_1 \, da_1 + \int_{s_2} \bar{D}_2 \cdot \hat{n}_2 \, da_2 + \int_{s_3} \bar{D}_1 \cdot \hat{n}_3 \, da_3 + \int_{s_4} \bar{D}_2 \cdot \hat{n}_4 \, da_4.$$

If the ‘height’ of the Gaussian volume goes to zero (that is,  $\ell \rightarrow 0$ ), then  $\bar{D}_1 \cdot \hat{n}_3 \rightarrow 0$  and  $\bar{D}_2 \cdot \hat{n}_1 \rightarrow 0$ . Furthermore, if the area of the Gaussian volume is vanishingly small,  $\bar{D}$  will

not vary over the surface. Thus,

$$\int_v \vec{\nabla} \cdot \vec{D} dv = \vec{D}_1 \cdot \hat{n}_1 \int_{s_1} da_1 + \vec{D}_2 \cdot \hat{n}_2 \int_{s_2} da_2 = (\vec{D}_1 \cdot \hat{n}_1 + \vec{D}_2 \cdot \hat{n}_2) A = 0$$

with  $A$  = the area of the surface and  $S_1 = S_2$ . Combining the normal vectors  $\hat{n}_1 = -\hat{n}$  and  $\hat{n}_2 = \hat{n}$ , the above argument yields

$$(\vec{D}_2 - \vec{D}_1) \cdot \hat{n} = 0. \quad (6.1)$$

This is the first boundary condition. By analogy, utilizing  $\vec{\nabla} \cdot \vec{B} = 0$ , the second boundary condition is found to be

$$(\vec{B}_2 - \vec{B}_1) \cdot \hat{n} = 0. \quad (6.2)$$

The remaining two boundary conditions are found by considering the fields across a boundary enclosed by a Stokesian loop. Such a loop is pictured in Figure A.2. The closed contour is taken over  $C = C_1 + C_2 + C_3 + C_4$ . If  $\ell$ , the distance along the loop perpendicular to the boundary, is vanishingly small, then contours  $C_3$  and  $C_4$  are negligible. The integral around the loop is

$$\oint_C \vec{E} \cdot d\vec{\ell}.$$

The integral over the entire curve,  $C = C_1 + C_2$ , is

$$\oint_C \vec{E} \cdot d\vec{\ell} = \oint_{C_1} \vec{E}_1 \cdot d\vec{\ell}_1 + \oint_{C_2} \vec{E}_2 \cdot d\vec{\ell}_2.$$

Let the length of each part of this curve be  $w$ . When the electric field  $E$  is considered small over the curve  $C$ , then it is apparent that

$$\oint_C \vec{E} \cdot d\vec{\ell} = \vec{E}_1 \cdot (\hat{t} \times \hat{n}) w - \vec{E}_2 \cdot (\hat{t} \times \hat{n}) w.$$

When terms are combined the integral about the loop becomes

$$\oint_C \vec{E} \cdot d\vec{\ell} = (\vec{E}_1 - \vec{E}_2) \cdot (\hat{t} \times \hat{n}) w. \quad (6.3)$$

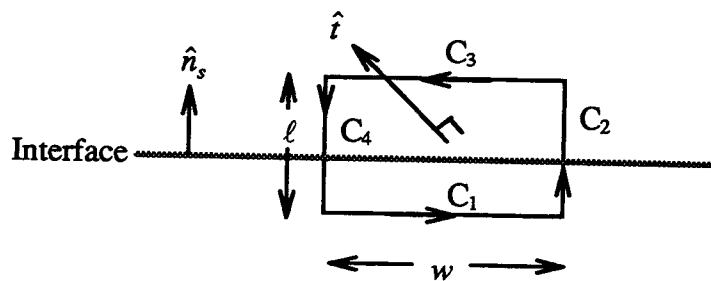


Figure A.2: A Stokesian loop straddling an interface.

The curl of  $E$ ,

$$\bar{\nabla} \times \bar{E} = -\frac{1}{c} \partial_t \bar{B},$$

when substituted into (6.3), yields

$$\int_s (\bar{\nabla} \times \bar{E}) \cdot \hat{n} \, da = \int_s -\frac{1}{c} (\partial_t \bar{B}) \cdot \hat{n} \, da = \oint_c \bar{E} \cdot d\bar{\ell}. \quad (6.4)$$

(6.4)

By Stoke's Theorem it is seen that reduces to

$$(\bar{E}_1 - \bar{E}_2) \cdot (\hat{t} \times \hat{n}) \, w = -\frac{1}{c} \int_s (\partial_t \bar{B}) \cdot \hat{n} \, da. \quad (6.5)$$

Since  $(\partial_t \bar{B})$  is finite over the surface  $S$  formed by the contour and  $\ell \rightarrow 0$ , the integral in (6.5) goes to zero, leaving

$$(\bar{E}_1 - \bar{E}_2) \cdot (\bar{t} \times \hat{n}) = 0$$

or

$$[(\bar{E}_1 - \bar{E}_2) \cdot \bar{t}] \times \hat{n} = 0. \quad (6.6)$$

This is the statement that the tangential components of  $E$  are continuous. By analogy the fourth boundary condition is derived:

$$\left[ (\vec{H}_1 - \vec{H}_2) \cdot \vec{t} \right] \times \hat{n} = 0. \quad (6.7)$$

The wave equation is derived in the usual way by considering the curl of the curl of  $E$ :

$$\vec{\nabla} \times \vec{\nabla} \times \vec{E} = -\frac{1}{c} \vec{\nabla} \times (\partial_t \vec{B}). \quad (6.8)$$

Making no assumptions about the form of the permittivity,  $\epsilon$ , and permeability,  $\mu$ , equation (6.8) yields

$$\vec{\nabla}(\vec{\nabla} \cdot \vec{E}) - \nabla^2 \vec{E} = -\frac{1}{c} \partial_t (\vec{\nabla} \times \vec{B}). \quad (6.9)$$

Recalling  $\vec{\nabla} \cdot \vec{D} = 0$  and  $\vec{D} = \epsilon \vec{E}$ , it is then realized that

$$\vec{\nabla} \cdot \vec{D} = \vec{\nabla} \cdot (\epsilon \vec{E}) = \vec{E} \cdot \vec{\nabla} \epsilon + \epsilon \vec{\nabla} \cdot \vec{E} = 0.$$

By rearranging this equation, an expression for the divergence of  $E$  can be found:

$$\vec{\nabla} \cdot \vec{E} = (\vec{E} \cdot \vec{\nabla} \epsilon) \epsilon^{-1}.$$

The first term in (6.9) can now be written as

$$\bar{\nabla}(\bar{\nabla} \cdot \bar{E}) = \frac{(\bar{E} \cdot \bar{\nabla} \epsilon) + (\bar{\nabla} \epsilon \cdot \bar{\nabla}) \bar{E} + \bar{\nabla} \epsilon \times (\bar{\nabla} \times \bar{E}) + \epsilon (\bar{E} \cdot \bar{\nabla} \epsilon) \bar{\nabla} \left( \frac{1}{\epsilon} \right)}{\epsilon}. \quad (6.10)$$

Recall that  $\bar{B} = \mu \bar{H}$ . Since the curl of  $H$  is given by

$$\bar{\nabla} \times \bar{H} = \bar{\nabla} \times \left( \frac{\bar{B}}{\mu} \right) = \frac{1}{c} \partial_t \bar{D} = \frac{1}{c} \partial_t (\epsilon \bar{E}),$$

then, applying the relationship between  $B$  and  $H$ , it is clear that

$$\bar{\nabla} \times \bar{B} + \mu \left( \bar{\nabla} \frac{1}{\mu} \right) \times \bar{B} = \frac{(\epsilon \mu)}{c} \partial_t \bar{E},$$

where  $\partial_t \epsilon = 0$ . Herein lies the first assumption, that the permittivities of the waveguide materials, and hence their indices of refraction, are not time dependent (at least on the scale of the oscillations of the wave). Thus, the curl of  $B$  becomes

$$\bar{\nabla} \times \bar{B} = \frac{n^2}{c} \partial_t \bar{E} - \mu \left( \bar{\nabla} \frac{1}{\mu} \right) \times \bar{B}. \quad (6.11)$$

Combining (6.10), (6.9), and (6.11); gives the most general form for the wave equation,

$$\nabla^2 \bar{E} - \frac{n^2}{c^2} \partial_t^2 \bar{E} = \mu \left( \bar{\nabla} \frac{1}{\mu} \right) \times (\bar{\nabla} \times \bar{E})$$

$$= \frac{(\bar{E} \cdot \bar{\nabla} \epsilon) + (\bar{\nabla} \epsilon \cdot \bar{\nabla}) \bar{E} + \bar{\nabla} \epsilon \times (\bar{\nabla} \times \bar{E}) + \epsilon (\bar{E} \cdot \bar{\nabla} \epsilon) \bar{\nabla} \left( \frac{1}{\epsilon} \right)}{\epsilon}$$

and, by analogy, it can be shown that

$$\nabla^2 \bar{B} - \frac{n^2}{c^2} \partial_t^2 \bar{B}$$

$$= -\mu \left( \bar{\nabla} \frac{1}{\mu} \cdot \bar{\nabla} \right) \bar{B} - \mu \bar{B} \nabla^2 \frac{1}{\mu} + \mu (\bar{B} \cdot \bar{\nabla}) \left( \bar{\nabla} \frac{1}{\mu} \right) + \mu \left[ \left( \bar{\nabla} \frac{1}{\mu} \right) \times (\bar{\nabla} \times \bar{B}) \right]$$

$$- \frac{\mu}{n^2} [(\bar{\nabla} \epsilon) \times (\bar{\nabla} \times \bar{B})] - \frac{\mu^2}{n^2} \left[ (\bar{\nabla} \epsilon) \times \left( \bar{\nabla} \frac{1}{\mu} \right) \times \bar{B} \right].$$

In a dielectric medium  $\mu$  is constant, therefor  $\bar{\nabla} \mu^{-1} = 0$ . For a dielectric with a possible index gradient, the wave equations are

$$\nabla^2 \bar{E} - \frac{n^2}{c^2} \partial_t^2 \bar{E}$$

$$= - \frac{2(\bar{E} \cdot \bar{\nabla} n) \bar{\nabla} n + 2(\bar{\nabla} n \cdot \bar{\nabla}) \bar{E} + 2 \bar{\nabla} n \times (\bar{\nabla} \times \bar{E}) + n^2 (\bar{E} \cdot \bar{\nabla} n) \bar{\nabla} \left( \frac{1}{n^2} \right)}{n^2}$$

(6.12)

and

$$\nabla^2 \bar{B} - \frac{n^2}{c^2} \partial_t^2 \bar{B} = \frac{2}{n^2} [(\bar{\nabla} n) \times (\bar{\nabla} \times \bar{B})] \quad (6.13)$$

The wave equations for a waveguide, (6.12) and (6.13), are perfectly general for dielectric waveguides. This includes graded-index and step index waveguides (i.e., planar structures for which  $n = n(r)$ ). Solutions to these inhomogeneous differential equations can be solved by adding a particular solution to the solution of the homogenous wave equation. Another approach, treating the modes as quantized solutions to the Hamiltonian of the situation and engaging perturbation theory, is discussed in the main text.

Given the planar geometry, it is evident that two separate polarizations (of the electric field) are possible. That is, transverse waves occupy the modes of these waveguides. These polarizations are referred to as transverse electric (TE), when the electric field in the plane of the waveguide. In the other polarization, transverse magnetic (TM), the magnetic field is in the plane of the waveguide (along the y-axis). Separating the components of Maxwell's equations, two sets of equations are found. For TE polarizations ( $E_z = E_x = H_y = 0$ ) the following expressions hold:

$$\partial_z E_x = 0, \quad \partial_x E_z = 0, \quad \partial_y H_x = 0, \quad \partial_y H_z = 0, \quad B_y = 0, \quad \partial_t B_y = 0,$$

$$\partial_z E_y = \frac{1}{c} \partial_t B_x, \quad \text{and} \quad \partial_x H_z - \partial_z H_x = \frac{1}{c} \partial_t D_y.$$

Therefore, the wave equations for TE polarizations reduce to

$$E_x = 0,$$



$$\begin{aligned} & \partial_x^2 E_y + \partial_y^2 E_y + \partial_z^2 E_y - \frac{n^2}{c^2} \partial_t^2 E_y \\ &= -\frac{2}{n^2} \left[ E_y \partial_y^2 n + (\partial_x n \partial_x + \partial_y n \partial_y + \partial_z n \partial_z) E_y \right] \\ & \quad + \frac{n^2}{2} E_y (\partial_y n) \partial_y \left( \frac{1}{n^2} \right) \end{aligned}$$

$$E_z = 0,$$

$$\partial_x^2 B_x + \partial_z^2 B_x + \partial_z^2 B_x - \frac{n^2}{c^2} \partial_t^2 B_x + \frac{1}{n^2} (\partial_z B_x - \partial_x B_z) \partial_z n^2 = 0,$$

$$\partial_y B_x = 0,$$

$$B_y = 0,$$

$$\partial_x^2 B_z + \partial_z^2 B_z + \partial_z^2 B_z - \frac{n^2}{c^2} \partial_t^2 B_z + \frac{1}{n^2} (\partial_x B_z - \partial_z B_x) \partial_x n^2 = 0,$$

and

$$\partial_y B_z = 0.$$

For TM polarizations ( $E_y = H_z = H_x = 0$ ), the components of Maxwell's equations appear as:

$$\partial_y E_z = 0, \quad \partial_y E_x = 0, \quad D_y = 0, \quad \partial_x H_z = 0, \quad \partial_z H_x = 0, \quad \partial_t D_y = 0,$$

$$\partial_x E_z - \partial_z E_x = -\frac{1}{c} \partial_t B_y, \quad \partial_z H_y = -\frac{1}{c} \partial_t D_x, \quad \text{and} \quad \partial_x H_y = \frac{1}{c} \partial_t D_z.$$

The wave equations for TM polarizations therefor become

$$\partial_x^2 E_x + \partial_z^2 E_x - \frac{n^2}{c^2} \partial_t^2 E_x = 0,$$

$$\partial_y E_x = 0,$$

$$E_y = 0,$$

$$\partial_x^2 E_z + \partial_z^2 E_z - \frac{n^2}{c^2} \partial_t^2 E_z = 0,$$

$$\partial_y E_z = 0,$$

$$B_x = 0,$$

$$(\partial_x B_y) \partial_y n^2 = 0,$$

$$\partial_x^2 B_y + \partial_z^2 B_y + \partial_t^2 B_y - \frac{n^2}{c^2} \partial_t^2 B_y + \frac{1}{n^2} \left( (\partial_z B_y) (\partial_z n^2) + (\partial_x B_y) (\partial_x n^2) \right) = 0,$$

$$B_z = 0,$$

and

$$(\partial_z B_y) \partial_y n^2 = 0.$$

These expressions are general; the only assumption made is that the waveguides are made of dielectric materials. In particular that  $\mu$  is unity in all the media comprising the waveguide. However, solutions for graded-index waveguides are not easy to come by.

For step index waveguides, or 'simple waveguides', (6.12) and (6.13) become homogeneous and solutions are straightforward to find. The next section discusses the solutions to asymmetric step index waveguides.

## A.2 Field Solutions for Asymmetric Step Index Dielectric Waveguides

The wave equations for a waveguide given in (6.12) and (6.13) were derived for dielectric waveguides in a perfectly general way. This was done so as to include the possibility of graded-index waveguides. However, the simplest waveguide is one without index gradients, that is when  $\bar{\nabla}n = 0$  except at a boundary. This is the case solved here. Solutions to the wave equation are found in the usual way for a general step index waveguide. Here, a brief review is given with the emphasis placed on the asymmetric nature of the waveguides. Define a primed coordinate system that is rotated about the  $y'$ -axis an angle  $-\vartheta$  and then about the  $x'$ -axis by an angle  $-\alpha$  from the unprimed system in Figure A.3.

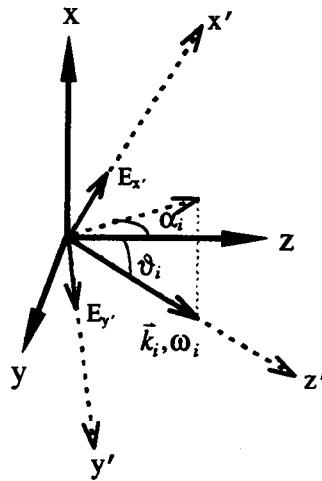


Figure A.3: One of the beam coordinate systems and the (unprimed) waveguide coordinate system.

Consider a wave traveling in  $z'$ -direction; any general polarization (direction of the electric field) can be written as a linear combination of two polarizations in the  $\hat{x}'$ - and  $\hat{y}'$ -directions:

$$\vec{E}' = E_{x'}\hat{x}' + E_{y'}\hat{y}'.$$

The general polarization can also be written in terms of a superposition of TE and TM polarizations by defining  $\vec{E}'_{TE} \equiv E_{y'}\hat{y}'$  and  $\vec{E}'_{TM} \equiv E_{x'}\hat{x}'$ . Then, any general polarization state can be represented as  $\vec{E}' = \vec{E}'_{TE} + \vec{E}'_{TM}$ . In the unprimed coordinate system, the field can be represented as

$$\vec{E} = E_x\hat{x} + E_y\hat{y} + E_z\hat{z}$$

or

$$\begin{aligned} \vec{E} = E_{x'} \cos\vartheta \hat{x} + E_{x'} \sin\vartheta \sin\alpha \hat{y} + E_{x'} \sin\vartheta \cos\alpha \hat{z} + \\ E_{y'} \cos\alpha \hat{y} - E_{y'} \sin\alpha \hat{z} \end{aligned}$$

In terms of the TE polarization in the unprimed frame, the fields are described by

$$\vec{E}_{TE} = E_{y'} \cos\alpha \hat{y} - E_{y'} \sin\alpha \hat{z} \quad (6.14)$$

and in terms of the TM polarization,

$$\vec{E}_{TM} = E_{x'} \cos\vartheta \hat{x} + E_{x'} \sin\vartheta \sin\alpha \hat{y} + E_{x'} \sin\vartheta \cos\alpha \hat{z}. \quad (6.15)$$

With expressions (6.14) and (6.15) and the solutions to the wave equations, any combination of waves (e.g., Gaussian beams) can be described. Note that the wavevector in the unprimed frame is

$$\vec{k} = -k_z \sin \vartheta \hat{x} + k_z \cos \vartheta \sin \alpha \hat{y} + k_z \cos \vartheta \cos \alpha \hat{z}, \quad (6.16)$$

with

$$|\vec{k}| = k = nk_o = n \frac{\omega}{c}.$$

Here  $k_o$  is related to the vacuum wavelength,  $\lambda_o$ , by  $k_o = 2\pi/\lambda_o$ . This framework allows for a relation between the field strength of the waves (or beams) outside and inside the waveguide to be derived.

An example of an asymmetric waveguide is shown in Figure 2.4. There, the index of refraction changes along the perpendicular, or  $x$ -direction, in a discontinuous way. Recall that the coordinate system places the origin in the center of the waveguide. The origin of this coordinate system becomes the center of the interaction volume when phase-matching issues are examined. If the film thickness is  $h$ , then the cover begins at  $h/2$  and the substrate-film interface is found at  $-h/2$ . Thus, the index dependence on  $x$  could be written as:

$$n = n(x) = \begin{cases} n_c & \text{for } x > \frac{h}{2} \\ n_f & \text{for } |x| < \frac{h}{2} \\ n_s & \text{for } x < -\frac{h}{2}, \end{cases} \quad (6.17)$$

where  $n_c$  is the index of the cover,  $n_f$  is the index of the film, and  $n_s$  is the index of the substrate. It is assumed that infinite plane waves with TE polarization are impinging at the film-cover boundary of the waveguide. Suppose the electric field at the cover-film boundary has the form

$$\bar{E}(x, y, z, t) = E_y \hat{y} e^{i(\bar{k} \cdot \bar{z} - \omega t)}, \quad (6.18)$$

where

$$\bar{k} \cdot \bar{z} = k_0 n_f z \sin \theta. \quad (6.19)$$

A mode-related wavevector for the film is defined as

$$\beta_v \equiv k_0 n_f \sin \theta, \quad (6.20)$$

where  $\theta$  is the angle with which the waves are impinging on the surface as measured in the film (as presented in (6.18)). The angle  $\vartheta$  *outside* the waveguide can be measured by the experimenter as presented in (6.16). The electric field inside the waveguide can be expressed as

$$\bar{E}(x, y, z, t) = E_y f(x) e^{i\beta_v z} e^{-i\omega t}, \quad (6.21)$$

where  $f(x)$  is some function such that  $f(0) = 1$  in order to satisfy the initial condition that

the field satisfies (6.18). Applying this form of the electric field to the wave equation, expression (6.12), gives the Helmholtz equation:

$$\partial_x^2 f - \beta_v^2 + \frac{n^2}{c^2} \omega^2 f = 0. \quad (6.22)$$

Defining an effective wavevector as

$$K^2 \equiv \frac{n^2}{c^2} \omega^2 - \beta_v^2, \quad (6.23)$$

the Helmholtz equation (6.22) becomes

$$\partial_x^2 f + K^2 f = 0. \quad (6.24)$$

This differential equation, equation (6.24), has solutions

$$f(x) = \begin{cases} e^{\pm iKx} & \text{for } K > 0 \\ e^{\pm Kx} & \text{for } K < 0. \end{cases} \quad (6.25)$$

Definitions of the electric field amplitude in each region of the waveguide can be made such that

$$E_c \equiv E_y(x > h/2) \quad \text{for the cover,} \quad (6.26)$$

$$E_f \equiv E_y(|x| < h/2) \quad \text{for the film, and} \quad (6.27)$$

$$E_s \equiv E_y(x < -h/2) \text{ for the substrate.} \quad (6.28)$$

Note that these amplitudes are written in terms of the amplitude of the impinging wave as in (6.18). Similarly, effective wavevectors  $K_c$ ,  $K_f$ , and  $K_s$  can be defined for these three regions as

$$K_c^2 \equiv k^2 n_c^2 - \beta_v^2 \text{ for the cover,} \quad (6.29)$$

$$K_f^2 \equiv -k^2 n_f^2 + \beta_v^2 \text{ for the film, and} \quad (6.30)$$

$$K_s^2 \equiv k^2 n_s^2 - \beta_v^2 \text{ for the substrate.} \quad (6.31)$$

The amplitudes of the fields in the three regions can then be written as

$$E_c = ae^{-K_c(x-h/2)}, \quad (6.32)$$

$$E_f = f \cos(K_f x - \phi), \quad (6.33)$$

and

$$E_s = be^{-K_s(h/2-x)}. \quad (6.34)$$

Here,  $\phi$  is the phase of the wave in the film is yet undetermined. This phase is found by considering the boundary conditions. Applying the boundary conditions leads to

$$\frac{K_c}{K_f} = \tan\left(K_f \frac{h}{2} - \phi\right) = \tan\left(K_f \frac{h}{2} - \phi - m\pi\right) \quad (6.35)$$



and

$$\frac{K_s}{K_f} = \tan\left(K_f \frac{h}{2} + \phi'\right) = \tan\left(K_f \frac{h}{2} + \phi' - m'\pi\right), \quad (6.36)$$

with  $m$  and  $m'$  integers. Definitions for the phases of the oscillating fields in the cover and substrate can be made such that

$$\phi_s \equiv \tan^{-1}\left(\frac{K_s}{K_f}\right), \quad (6.37)$$

and

$$\phi_c \equiv \tan^{-1}\left(\frac{K_c}{K_f}\right). \quad (6.38)$$

Thus, the added phases of the waves in (6.35) and (6.36) are determined as

$$\phi' = -K_f\left(\frac{h}{2}\right) + \phi_s + m'\pi, \quad (6.39)$$

and

$$\phi = K_f\left(\frac{h}{2}\right) - \phi_c - m\pi. \quad (6.40)$$

Since  $\phi' = \phi$ , expressions (6.39) and (6.40) combine to give

$$K_f h - \phi_s - \phi_c = (m + m')\pi \equiv v\pi. \quad (6.41)$$

Recall that  $K_f$  depends on the effective wavevector  $\beta_v$ . Equation (6.41) indicates that the effective wavevectors  $\beta_v$  are discrete. Evidently, these discrete wavevectors describe the allowed modes a waveguide. The 'new' propagation constant  $v$  is the mode identifier; it runs over all positive integers.

The normalization constants  $a$ ,  $b$ , and  $f$  in expressions (6.32), (6.33), and (6.34) can be found in the usual way, by squaring the solutions to  $f(x)$ . The results give the final form for the electric field in the waveguide:

$$E_y = \begin{cases} \left[ \frac{(n_f^2 - N_v^2)}{(n_f^2 - n_c^2)} \right]^{-\frac{1}{2}} e^{-K_c(x-h/2)} e^{i\beta_v z} e^{-i\omega t} & \text{for } x > h/2 \\ \cos(K_f x - \phi_v) e^{-i\beta_v z} e^{-i\omega t} & \text{for } |x| < h/2 \\ \left[ \frac{(n_f^2 - N_v^2)}{(n_f^2 - n_s^2)} \right]^{-\frac{1}{2}} e^{+K_s(h/2+x)} e^{i\beta_v z} e^{-i\omega t} & \text{for } x < -h/2, \end{cases} \quad (6.42)$$

with the following definitions

$$N_v \equiv \beta_v / k, \quad (6.43)$$

$$\phi_v \equiv K_f \frac{h}{2} - \phi_c = -K_f \frac{h}{2} + \phi_s, \quad (6.44)$$

$$v\pi = K_f h - \phi_c - \phi_s, \quad (6.45)$$

and

$$\beta_v \equiv kn_f \sin \theta_v. \quad (6.46)$$

To find the allowed angles  $\theta_v$ , a transcendental equation must be solved:

$$\tan\left(\left[K_f\right]_v h - v\pi\right) = \left[K_f \frac{K_c + K_s}{K_f^2 - K_c K_s}\right]_v. \quad (6.47)$$

These final expressions describe the field amplitude and phase of a guided wave of frequency  $\omega$  for all three regions of the asymmetric step index waveguide. Looking at (6.42), in the film there is a traveling wave with an effective wavevector that depends on a discrete mode of the waveguide. In the cover and substrate regions, there is an exponentially decaying evanescent field. The depth to which these fields penetrate the region also is mode dependent. Expression (6.42) describes a guided wave of a single frequency. To accommodate four-wave mixing experiments in a waveguide, four guided waves are required. The next section discusses the four-wave mixing process in a waveguide, determines the generated intensity, and derives the ratios used to quantify the surface (i.e., cover) specificity.

### A.3 The Intensity, Phase-Mismatch, and Quantitative Tools for WISOM

Four-wave mixing in step index waveguides can be surface-specific if the induced polarization due to the incident waves is (nearly) zero and non-zero in the cover region. In this treatment, the intensity contribution from the cover and film (including substrate) are found separately. To quantify the surface-specificity, two ratios are introduced which are sensitive to the cover-to-film ratio. As stated in the main text, the cover-to-film ratio diverges as the film contribution goes to zero. A more useful set of interpretive tools are derived. One is sensitive to the amplitude of the cover contribution relative to the overall signal, while the other is also sensitive to the phase of the cover contribution in relation to the total contribution. It is also important to know the phase-mismatch between the induced polarization and the generated (mixed) wave. Only infinite plane waves are considered in this treatment.

Consider that the four waves of a FWM process have wavevectors in the film region which can be defined as

$$k_1 \equiv n(\omega_1) \frac{\omega_1}{c}, \quad k_2 \equiv n(\omega_2) \frac{\omega_2}{c}, \quad k_3 \equiv n(\omega_3) \frac{\omega_3}{c}, \quad \text{and} \quad k_4 \equiv n(\omega_4) \frac{\omega_4}{c}. \quad (6.48)$$

Phase-mismatching is, in general, given by

$$\Phi(x, y, z) \equiv \Delta \vec{k} \cdot \vec{r} \quad (6.49)$$

with

$$\Delta \vec{k} \equiv (\vec{k}_1 + \vec{k}_2 + \vec{k}_3) + \vec{k}_4, \quad (6.50)$$

where  $\vec{k}_4$  is the wavevector of the observed photon. Again, the convention is that positive wavevectors indicate annihilation, negative wavevectors indicate creation of photons the OM process. It is understood that since  $\vec{k}_4$  is observed, this wavevector is negative. The phase-mismatch for the possible guided modes  $v$  is

$$\Phi = \Delta \vec{k}_v \cdot \hat{z} \frac{\ell}{2}, \quad (6.51)$$

where

$$\begin{aligned} \Delta \vec{k}_v \cdot \hat{z} = & k_1 n(\vec{r}) \sin \theta_{v_1} \cos \alpha_{k_1} + \\ & k_2 n(\vec{r}) \sin \theta_{v_2} \cos \alpha_{k_2} + \\ & k_3 n(\vec{r}) \sin \theta_{v_3} \cos \alpha_{k_3} + \\ & k_4 n(\vec{r}) \sin \theta_{v_4} \cos \alpha_{k_4} \end{aligned} \quad (6.52)$$

and  $\ell$  is the length over which the three incident waves overlap. Of course for plane waves  $\ell$  is infinite, but plane wave are not used in practice. Thus,  $\ell$  is approximated by some experimentally reasonable value. Recall the angles  $\alpha$  are the angles in the plane of the waveguide that each wave propagates at. The famous result for the intensity of the

fourth wave generated in a FWM process appears as

$$I_{\omega_4} = \left| \int \int \int_{\substack{\text{all space} \\ \text{all time}}} \bar{E}_4 \cdot \chi^{(3)} : \bar{E}_1 \bar{E}_2 \bar{E}_3 \frac{\sin \Phi}{\Phi} dx dy dz dt \right|^2 \quad (6.53)$$

Replacing the electric fields with the form defined in (6.21), the intensity of the mixed wave becomes

$$I(\omega_4) = \left( \frac{2\pi}{k_4} \right)^2 \left( \frac{\omega_4}{c} \right)^4 I_{\omega_1} I_{\omega_2} I_{\omega_3} \times \left| \int_{-\infty}^{\infty} \hat{\epsilon}_4 \cdot \chi^{(3)} : \hat{\epsilon}_1 \hat{\epsilon}_2 \hat{\epsilon}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \right|^2, \quad (6.54)$$

where  $I_{\omega_1}$ ,  $I_{\omega_2}$ , and  $I_{\omega_3}$  are the maximum amplitudes of the incident waves at  $\omega_1$ ,  $\omega_2$ , and  $\omega_3$ , respectively. The estimated interaction length is  $\ell$ . The integral can be separated into two parts: the integral over the cover region and the integral over the film and substrate regions. Typically the nonlinear susceptibilities are different in this region. This is indicated by this form for the susceptibility:

$$\chi^{(3)}(x) = \begin{cases} \chi_c^{(3)} & \text{for } x > \frac{h}{2} \\ \chi_f^{(3)} & \text{for } |x| < \frac{h}{2} \\ \chi_s^{(3)} & \text{for } x < -\frac{h}{2}. \end{cases} \quad (6.55)$$

Thus (6.54) can be rewritten as

$$\begin{aligned}
 I(\omega_4) = & \left( \frac{2\pi}{k_4} \right)^2 \left( \frac{\omega_4}{c} \right)^4 I_{\omega_1} I_{\omega_2} I_{\omega_3} \\
 & \times \left| \begin{aligned} & \int_{-\infty}^{-h/2} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_s^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \\ & + \int_{-h/2}^{+h/2} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_f^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \\ & + \int_{+h/2}^{+\infty} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_c^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \end{aligned} \right|^2 .
 \end{aligned} \tag{6.56}$$

Consider that the first two terms in (6.56) are the ‘background’ portion,  $E_B$ , of the generate field and the last term is the cover-specific anecdotal field,  $E_A$ :

$$E_A \equiv \int_{+h/2}^{+\infty} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_c^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \tag{6.57}$$

and

$$\begin{aligned}
 E_B \equiv & \int_{-\infty}^{-h/2} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_s^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx \\
 & + \int_{-h/2}^{+h/2} \hat{\mathbf{e}}_4 \cdot \boldsymbol{\chi}_f^{(3)} : \hat{\mathbf{e}}_1 \hat{\mathbf{e}}_2 \hat{\mathbf{e}}_3 f_1(x) f_2(x) f_3(x) \ell(x) \operatorname{sinc} \Phi(x) dx .
 \end{aligned} \tag{6.58}$$

These definitions make the total intensity proportional to the square of the sum of the background and anecdotal fields:

$$I(\omega_4) \sim |E_A + E_B|^2 = |E_A|^2 + |E_B|^2 + E_A^* E_B + E_A E_B^*. \quad (6.59)$$

The direct ration (DR) and cross ratio (CR) can be defined as

$$DR \equiv \frac{|E_A|^2}{|E_A|^2 + |E_B|^2} \quad (6.60)$$

and

$$CR \equiv \frac{E_A^* E_B + E_A E_B^*}{|E_B|^2}. \quad (6.61)$$

The denominator of the DR does not include the cross terms that the total intensity does. If the fields from the surface and film are anti-parallel and of the same magnitude, the total intensity is zero. Under these conditions, the DR remains finite. A DR value of one indicates excellent surface specificity. The DR is insensitive to the phases of the fields. In particular, when the bulk and cover generated fields are of the same magnitude but anti-parallel, the DR is zero indicating a poor signal-to-noise ratio. This is clearly misleading. Hence the CR is introduced to provide a measure of the sizes of the cross terms in the intensity. A large positive value of the CR indicates parallel fields where the magnitude of the surface is much greater than the bulk whereas a large negative CR indicates parallel fields with the cover field much larger than the bulk-generated field.



## APPENDIX B

This appendix is devoted to determining the solutions for the generated fields due to a nonlinear source term. Specifically, expressions for the fields ‘reflected’ from a layered structure of nonlinear media as described in Chapter 3 are sought. There are three essential steps in determining the generated nonlinear fields at a surface. First, the amplitudes of these fields at a boundary between two media at least one of which is nonlinear must be found. The treatment provided here was assembled from Bloembergen’s work. Since a portion of the method parallels the derivation of the linear Fresnel equations, the amplitude coefficients are called nonlinear Fresnel coefficients. Inside a nonlinear medium, the amplitude of the generated wave increases according to the distance the wave traverses in the medium and the phase-mismatch between the induced polarization and the generated wave. Determining this factor comprises the second step. Since this macroscopic model requires three interfaces (the vacuum-surface, the surface-bulk median, and the bulk median-bulk interfaces), the nonlinear Fresnel coefficients must be expressed in terms of the interfacial parameters. This final step is accomplished by matching the boundary conditions at each interface.

### **B.1 Determining the Solutions to the Inhomogeneous Wave Equation**

Bloembergen and co-workers determined the theoretical equations that govern optical mixing, particularly SHG, in reflection geometry. Bloembergen makes the assumption that ‘mixed waves’ (waves generated due to a nonlinear susceptibility) emanate or ‘reflect’ from the interface. These waves are produced by a source, or inhomogeneous wave; it is created by a nonlinear polarization induced by all incident fields. With this assumption, this inhomogeneous source wave is derived from the inhomogeneous wave equation in a nonlinear dielectric. Generalizations to the well-known laws of reflection and refraction

give the direction of propagation of the mixed waves. These laws can be considered generalized, nonlinear, Fresnel coefficients.

When fields  $i = 1$  through  $i = n - 1$  are applied to a medium, a nonlinear polarization is induced via the susceptibility. The induced polarization oscillates at the mixing frequency. The oscillating polarization will, in turn, radiate energy in the form of a traveling wave of the same frequency. Consider Maxwell's equations in Gaussian units with the displacement field

$$\vec{D} = \epsilon \vec{E} + 4\pi \vec{P}_{non-linear}$$

where, in general,

$$\vec{P}_{non-linear}(\vec{r}', t') = \sum_n \chi^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}) : \vec{E}(\omega_1) \dots \vec{E}(\omega_{n-1}) e^{i\vec{k}_s \cdot \vec{r}'} e^{-i\omega_n t'}.$$

The oscillating polarization which could be seen as a source wave. Note that this source wave acts as if it were in a medium with effective index  $n_s$ , defined as

$$n_s \equiv \frac{c}{\omega_n} |\vec{k}_s| = \frac{c}{\omega_n} \left| \sum_{i=1}^{n-1} \vec{k}_i(\omega_i) \right|, \quad (7.1)$$

Suppose that the electric field created by this oscillating polarization is observed at location  $(\vec{r}, t)$ . (Refer to Figure B.1.) Anticipating the use of Green's Theorem, the displacement field at  $(\vec{r}, t)$  can be written in terms of the displacement field at  $(\vec{r}', t')$ :

$$D(\vec{r}, t) = \int d\vec{r}' dt' G(\vec{r}, t; \vec{r}', t') D(\vec{r}', t') \quad (7.2)$$

The displacement field is used since there may be additional included polarization at  $(\vec{r}, t)$ ; its use simplifies bookkeeping. The Green's function must assume the form of a spherically expanding wave beginning at  $(\vec{r}', t')$  such that

$$G(\vec{r}, t; \vec{r}', t') = \frac{1}{|\vec{r} - \vec{r}'|} \delta\left(t' - \frac{t - |\vec{r} - \vec{r}'|}{c/n(\omega)}\right) \tilde{\nabla}_{\mu, \nu}, \quad (7.3)$$

where  $\tilde{\nabla}_{\mu, \nu}$  is a dyadic operator which resolves the components of the displacement field when  $\vec{P}_{non-linear}$  and  $\vec{k}_s$  are in different directions.

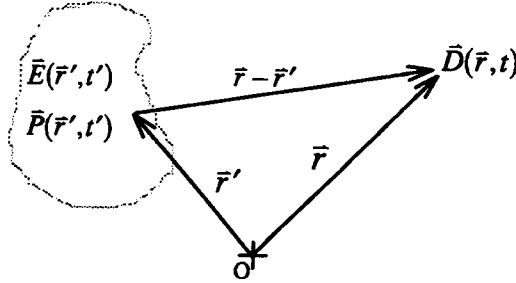


Figure B.1: A polarization and propagating field at  $(\vec{r}', t')$  and the displacement field observed at  $(\vec{r}, t)$ .

Assuming  $\mu = 1$ , the media are not conductors, and  $\partial_t^2 \epsilon = \partial_t \epsilon = 0$ , it follows that the wave equation is a second order inhomogeneous differential equation as described by

$$\tilde{\nabla} \times \tilde{\nabla} \times \vec{E}(\vec{r}, t) = -\frac{\mu}{c^2} \partial_t^2 \vec{D}(\vec{r}, t). \quad (7.4)$$

Here the expression  $\bar{B} = \mu\bar{H}$  has been invoked. Using  $\bar{D} = \epsilon\bar{E} + 4\pi\bar{P}$ , (7.4) becomes

$$\bar{\nabla} \times \bar{\nabla} \times [\bar{D}(\bar{r}, t) - 4\pi\bar{P}_{non-linear}(\bar{r}, t)] = -\frac{n_t^2}{c^2} \partial_t^2 \bar{D}(\bar{r}, t). \quad (7.5)$$

Equation (7.5) reduces to a Helmholtz-like equation with an inhomogeneous term:

$$\bar{\nabla}_{\mu,v} \bar{D}(\bar{r}, t) + \frac{n_t^2}{c^2} \partial_t^2 \bar{D}(\bar{r}, t) = 4\pi \bar{\nabla}_{\mu,v} \bar{P}_{non-linear}(\bar{r}, t), \quad (7.6)$$

where the dyadic operator  $\bar{\nabla}_{\mu,v}$  has been identified as  $\bar{\nabla}_{\mu,v} = \bar{\nabla} \times \bar{\nabla} \times \bar{V} = \bar{\nabla}(\bar{\nabla} \cdot \bar{V} - \nabla^2 \bar{V}) = \tilde{\epsilon}_{\mu,\alpha,v} \tilde{\epsilon}_{v,\beta,\gamma} \hat{e}_\mu \hat{e}_v \partial_{r_\alpha} \partial_{r_\beta} V_\gamma$ . A solution to (7.6) may be obtained by finding a general solution to the homogeneous case and then adding to it a specific solution of the inhomogeneous case. The homogeneous solutions are plane waves. The inhomogeneous solution is found by considering

$$\bar{\nabla}_{\mu,v} G(\bar{r}, t; \bar{r}', t') + \frac{n_t^2}{c^2} \partial_t^2 G(\bar{r}, t; \bar{r}', t') = -4\pi \delta(\bar{r} - \bar{r}') \delta(t - t'), \quad (7.7)$$

multiplying both sides through by  $d\bar{r}' dt'$   $G(\bar{r}, t; \bar{r}', t') D(\bar{r}', t')$ , and integrating over  $d\bar{r}'$  and  $dt'$ . The result is a specific form for the Green's function (used to find the  $\mu$ -th component of  $\bar{D}(\bar{r})$  arising from the  $v$ -th component of  $\bar{D}(\bar{r}')$ )

$$G(\bar{r}, \bar{r}'; \omega) = \delta(\bar{r} - \bar{r}') \left( \delta_{\mu v} - \frac{k_{s_\mu} k_{s_v}}{|\bar{k}_s|^2} \right) e^{i[\bar{k}_s \cdot (\bar{r} - \bar{r}')] } e^{-i\omega t}, \quad (7.8)$$

and the solution to the inhomogeneous equation

$$\bar{E}_{INHOMO}(\bar{r}, \omega_n) = \frac{-4\pi}{(n_t^2 - n_s^2)} \left( \bar{P}(\bar{r}, \omega_n) - \frac{\bar{k}_s (\bar{k}_s \cdot \bar{P}(\bar{r}, \omega_n))}{|\bar{k}_s|^2} \right) e^{i(\bar{k}_s \cdot \bar{r} - \omega_n t)}. \quad (7.9)$$

The above expression is found by evaluating (7.2) with (7.8), realizing that  $\bar{D}(\bar{r}) = n_t^2 \bar{E}(\bar{r}) + 4\pi \bar{P}(\bar{r})$ , but  $\bar{D}(\bar{r}') = n_s^2 \bar{E}(\bar{r}') + 4\pi \bar{P}(\bar{r}')$ . The index of refraction  $n_t$  is the true index of the medium at  $\bar{r}$ . Expression (7.9) is the one quoted by Bloembergen. One should note, however, that the build up in amplitude of the propagating wave as it encounters additional polarized media has been ignored. The final solution to the generated electric field due to a nonlinear polarization is found by adding this particular solution (7.9) to the homogeneous solution yielding:

$$\bar{E}_{total} = \bar{E}_{HOMO} + \bar{E}_{INHOMO} = \bar{E}_T \mathcal{E}_T e^{i(\bar{k}_T \cdot \bar{r} - \omega t)} + \bar{E}_S \mathcal{E}_S e^{i(\bar{k}_S \cdot \bar{r} - \omega t)}, \quad (7.10)$$

where  $\mathcal{E}_T$  must be determined by boundary conditions. The factor  $\bar{E}_S \mathcal{E}_S$  can be simplified by considering polarizations of the effective source wave.

As introduced in the main text, reflection geometry has two natural choices for the polarization basis vectors. As in linear reflection, the electric field can be decomposed into two directions, perpendicular ( $S$ ) and parallel ( $P$ ) to the plane of reflection. The plane of reflection is defined by the plane containing both the reflected and transmitted rays. Thus, the two polarizations can be decomposed:

$$\begin{aligned}
\bar{E}_{INHOMO}(\vec{r}, t) &= \bar{\epsilon}_S \mathcal{E}_{s,S} e^{i(\vec{k}_s \cdot \vec{r} - \omega_s t)} + \bar{\epsilon}_P \mathcal{E}_{s,P} e^{i(\vec{k}_s \cdot \vec{r} - \omega_s t)} \\
&= [\bar{E}_{INHOMO}(\vec{r}, t)]_S \bar{\epsilon}_S + [\bar{E}_{INHOMO}(\vec{r}, t)]_P \bar{\epsilon}_P.
\end{aligned} \tag{7.11}$$

The inhomogeneous fields with the two polarizations may be written

$$[\bar{E}_{INHOMO}(\vec{r}, t)]_S = \mathcal{E}_{s,S} e^{i(\vec{k}_s \cdot \vec{r} - \omega_s t)}$$

and

$$[\bar{E}_{INHOMO}(\vec{r}, t)]_P = \mathcal{E}_{s,P} e^{i(\vec{k}_s \cdot \vec{r} - \omega_s t)}.$$

The induced fields are, therefore, defined by (refer to (7.9))

$$\bar{\epsilon}_S \mathcal{E}_{s,S} \equiv \frac{-4\pi}{n_i^2 - n_s^2} \left( \bar{P}_S^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}) - \frac{\vec{k}_s (\vec{k}_s \cdot \bar{P}_S^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}))}{|\vec{k}_s|^2} \right) \tag{7.12}$$

and

$$\bar{\epsilon}_P \mathcal{E}_{s,P} \equiv \frac{-4\pi}{n_i^2 - n_s^2} \left( \bar{P}_P^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}) - \frac{\vec{k}_s (\vec{k}_s \cdot \bar{P}_P^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1}))}{|\vec{k}_s|^2} \right). \tag{7.13}$$

The susceptibility will be chosen to be scalar. This is done so that in the subsequent analysis the medium-independent angular ( $\theta$  and  $\phi$ ) dependence and phase-matching dependencies could be explored. In the case of  $S$ -polarization, there is no advantage to considering  $\vec{k}_s \cdot \vec{P}_P^{(n)} \neq 0$ . This is realized when considering a sheet of discrete radiating dipoles. Suppose the oscillating dipole moments are lined up and are in the  $S$ -plane. The dipole sheet radiation strength changes only with the *cosine of the polar angle*. There is no additional physics to be investigated. Thus, there is no need to consider any other case than  $\vec{k}_s \cdot \vec{P}_P^{(n)} \neq 0$ . On the other hand, for  $P$ -polarization,  $\vec{k}_s \cdot \vec{P}_P^{(n)} \neq 0$  is important to consider. Returning to the oscillating dipole sheet interpretation, when the dipole moments are in the  $P$ -plane, there can be no radiation in the direction of the induced polarization. When the process is linear (the dipoles oscillate with the same frequency as the applied fields), this condition occurs when the angle of reflection is Brewster's angle. By analogy, when the process is nonlinear, this condition occurs at the *nonlinear* Brewster's angle (NLBA). Thus, it is valuable to consider the angle (here called  $\alpha$ ) that the source wavevector makes with the induced polarization. The induced polarization is at an angle  $\beta$  with respect to the normal to the surface above. This situation is pictured in Figure B.2.

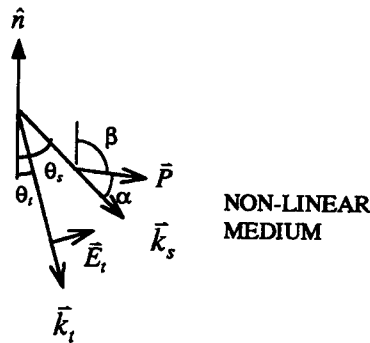


Figure B.2: Angle definitions for the nonlinear polarization, the source and generated wavevectors, and generated electric field.

With these cases in mind, the amplitudes of the inhomogeneous waves for  $S$  and  $P$  are

$$\mathcal{E}_{s,S} \equiv -4\pi \frac{|\tilde{P}_S^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1})|}{n_t^2 - n_s^2} \quad (7.14)$$

and

$$\mathcal{E}_{s,P} \equiv -4\pi \frac{|\tilde{P}_P^{(n)}(\omega_n; \omega_1, \dots, \omega_{n-1})|}{n_t^2 - n_s^2} \left( \cos\theta_s \sin\alpha + \sin\theta_s \cos\alpha \frac{(n_t^2 - n_s^2)}{n_t^2} \right), \quad (7.15)$$

where the angle  $\alpha = \pi - \beta - \theta_s$  is the angle between the polarization and the wavevector  $\vec{k}_s$ . I.e.,  $\vec{k}_s \cdot \vec{P}_P^{(n)} = |\vec{k}_s| |\vec{P}_P^{(n)}| \cos\alpha$ . The two terms in (7.15) are renderings of the vector products described in the parentheses of (7.13). Notice that the expression for the inhomogeneous source amplitudes for  $S$  (7.14) and  $P$  (7.15) polarization are identical when  $\theta_s = 0$  and polarization lies entirely in the plane of the surface (when  $\beta = \pi/2$ , i.e., when  $\alpha = \pi/2 - \theta_s = \pi/2$ ).

Note that in expressions (7.12) and (7.13), the local (Lorentz) field correction is *not* included, contrary to Bloembergen's suggestion. The expressions for the reflected fields including a bulk median are compared to Bloembergen's model (which does not include such a layer). Comparisons are made only with Bloembergen's expressions that do not contain the local field corrections.



## B.2 Explicitly Including the Phase-Mismatch Information Between Layers

The phase-mismatch information is paramount when comparing the phase-matched surface signal with the (possibly) phase-mismatched bulk signal. The approximations made in the Bloembergen model concerning phase-matched conditions are not appropriate when the phase-mismatch is large. The phase-mismatch is measured by the *phase-mismatch parameter*; it represents difference between the generated wavevector,  $\vec{k}_t$ , and the effective 'source' wavevector,  $\vec{k}_s$ . For instance, the argument  $\Phi_2$  is referred to as the *surface wave phase-mismatch parameter*. Similarly, the argument  $\Phi_3$  is the *bulk wave phase-mismatch parameter*. In simple situations, such as deep inside the bulk of a nonlinear crystal, a phase-mismatch parameter appears in the generated wave's phase factor as  $e^{-i\Phi}$ . As the wave progresses through the medium, it grows in intensity depending on the (square) of this phase factor integrated over the path length; this is the origin of the famous *Sinc* squared factor.

According to Bloembergen's theory, the integral over the region of overlap of the applied fields (commonly called the interaction region) is not necessary since the limit over  $d$  (the layer thickness) is taken. Taking this limit *before* performing the integration leaves an expression that neglects much of the phase-matching behaviour, specifically the *Sinc* behaviour of the field strength on the phase-mismatch parameter. It is unimportant in Bloembergen's analysis as phase-mismatched processes are never considered. When the phase-mismatch is large Bloembergen's expressions are inaccurate. If phase differences in the bulk and surface waves are to be investigated, the integration over the interaction volume (region) must be performed and must be an explicit part of the RISOM expressions. Inclusion of the integration in the RISOM expression not only repairs Bloembergen's result for the transmitted wave amplitudes (he neglects terms that are second order in the phase-mismatch for these amplitudes), but introduces an avenue to explore the phase-mismatch between surface and bulk waves.

Consider the strength of the *total* electric field in a nonlinear layer. At a particular location,  $r$ , it is the sum of the generated, homogeneous field,  $\bar{\epsilon}_T \mathcal{E}_T$ , and the inhomogeneous, source field  $\bar{\epsilon}_S \mathcal{E}_S$ :

$$\bar{E}_{total}(\bar{r}, t) = \bar{\epsilon}_T \mathcal{E}_T e^{i(\bar{k}_T \cdot \bar{r} - \omega t)} + \bar{\epsilon}_S \mathcal{E}_S e^{i(\bar{k}_S \cdot \bar{r} - \omega t)}. \quad (7.16)$$

Another way to write the total field is to consider the expression

$$\bar{E}_{total}(\bar{r}, t) = \bar{\epsilon}_{total} \mathcal{E}_{total} e^{i(\bar{k}_T \cdot \bar{r} - \omega t)}. \quad (7.17)$$

The total field inside the nonlinear medium can be transformed into the form of a single traveling wave, as in (7.17), by considering

$$\bar{E}_{total}(\bar{r}, t) = \left[ \bar{\epsilon}_T \mathcal{E}_T + \bar{\epsilon}_S \mathcal{E}_S e^{i[(\bar{k}_S - \bar{k}_T) \cdot \bar{r}]} \right] e^{i(\bar{k}_T \cdot \bar{r} - \omega t)}, \quad (7.18)$$

where the amplitude and polarization,  $\bar{\epsilon}_{total} \mathcal{E}_{total}$ , are given by the quantity in (7.18) inside the large square brackets. As the (total) wave propagates through the medium, it increases in field strength as  $r$  increases. The total field must begin with zero generated field, and grow in intensity as the inhomogeneous field adds strength to the homogeneous field. In a quantum mechanical sense, it refers to the stimulated emission of the field -- the field experiences gain. If a coordinate system is chosen (see Figure B.3) such that the wave propagates in the  $x$ - $z$ -plane,  $\bar{r} = [z / \cos \theta, \theta, 0]$ .

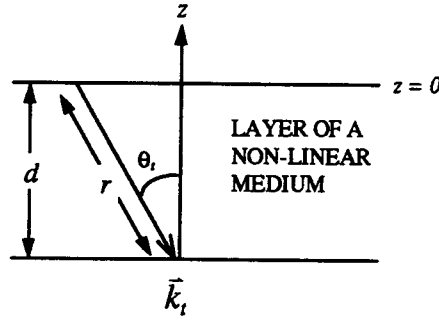


Figure B.3: Diagrammatic aid in determining the phase-mismatch in a nonlinear layer.

At  $z = 0$ , the surface of the medium, there is no total generated field. To express the coherent addition of the generated waves along the path  $\vec{r}$ , the integral over the inhomogeneous field must be taken. Thus, the strength of the total field a  $z$ -distance  $d$  from the surface is the integral over the source wave from zero to (total distance)  $r = d / \cos \theta_i$ :

$$\vec{E}_{total}(\vec{r}, \omega) = \frac{1}{d} \int_0^r d\vec{r}'' \vec{E}_S \mathcal{E}_S e^{i(\vec{k}_S - \vec{k}_T) \cdot \vec{r}''} e^{-i\omega t}. \quad (7.19)$$

With a change in variables of  $r' = r'' - r/2$ , the integral becomes

$$\int_{-r/2}^{r/2} d\vec{r}' e^{i(\vec{k}_S - \vec{k}_T) \cdot \vec{r}'} = d \operatorname{sinc}[(\vec{k}_S - \vec{k}_T) \cdot \hat{r} r / 2]$$

and realizing that  $\hat{r} = \hat{k}_T = \frac{\vec{k}_T}{|\vec{k}_T|}$  reduces to

$$\frac{d}{\cos \theta_T} \operatorname{sinc} \left[ - \left( |\vec{k}_T| - \frac{\vec{k}_S \cdot \vec{k}_T}{|\vec{k}_T|} \right) \frac{d}{2 \cos \theta_T} \right].$$

The phase-mismatch parameter may be defined as

$$\Phi \equiv - \left( |\bar{k}_T| - \frac{\bar{k}_S \cdot \bar{k}_T}{|\bar{k}_T|} \right) \frac{d}{2 \cos \theta_T}. \quad (7.20)$$

The total generated field in a nonlinear layer of thickness  $d$  is

$$\bar{E}_{total}(\bar{r}, t) = \bar{E}_S \mathcal{E}_S \sec \theta_T \text{sinc } \Phi e^{-i\Phi} e^{i(\bar{k}_T \cdot \bar{r} - \omega t)} \quad (7.21)$$

or 
$$\bar{E}_{total}(\bar{r}, t) = [\bar{E}_T \mathcal{E}_T + \bar{E}_S \mathcal{E}_S(\theta_T, \Phi)] e^{i(\bar{k}_T \cdot \bar{r} - \omega t)},$$

where  $\mathcal{E}_S(\theta_T, \Phi)$  has been defined. From (7.21) the generated field grows in intensity linearly with thickness as expected. It should be pointed out that when the phase-mismatch goes to zero, the *Sinc* is equal to one, and the expression reduces to the result of Bloembergen for small  $d$ .

### B.3 Determining the Electric Fields in the Layered Structure

Success in separating or enhancing the surface- over the bulk-generated waves lies in taking advantage of the geometry of the respective induced nonlinear polarizations. Recall that the induced nonlinear polarization is dependent on the nonlinear susceptibility (of surface or bulk regions) and the incident wave parameters. The incident wave parameters, namely their wavevectors (directions and frequencies) and polarizations, depend explicitly on the linear indices of refraction (and dispersions) of the media. In the model introduced in the main text of Chapter 3, any successful ‘experiment’ will be one in which the ratio of the surface-generated wave to the bulk median-generated wave amplitudes (as detected in the cover region) is favourable. The layered structure (see Figure 3.14) models the

structure of a surface, complete with bulk (including a nonlinear bulk median layer and a semi-infinite linear bulk region), through which optical mixing occurs. In this structure there are linear reflections governed by the linear Fresnel expressions and ‘nonlinear reflections’ governed by the expressions (7.14) through (7.15) given above. These expressions are also referred to as the nonlinear Fresnel expressions. Solutions can be found separately for the surface-generated and bulk median-generated signals as viewed in the cover region by imposing the boundary conditions on all the fields oscillating at the generated frequency (refer to the center and right portions of Figure 3.14). For the surface-generated wave, there are six unknown fields created from one known field. The ‘known’ field is the inhomogeneous wave which is dependent on the induced nonlinear polarization. Assuming a known surface-induced nonlinear polarization, solutions for all six fields, including the field in the cover region, can be found. Since there are two polarizations for the wave in the cover region, there are two sets of solutions, one for *S*- and one for *P*-polarization. The bulk-generated wave has similar solutions, obtained in the same manner. This and the following sections are devoted to finding these solutions.

The approach is to treat the six homogeneous field amplitudes as a vector, and the boundary conditions at each interface as a 6x6 matrix. The product of the homogeneous boundary condition matrix with the homogeneous field amplitude vector must yield a vector representing the satisfaction of the boundary conditions for the inhomogeneous waves. For instance, the homogeneous field amplitude vector for the surface region in *S*-polarization is  $V_{HOMO,surface,S}$  (see (7.22)). The homogeneous boundary condition matrix for the surface region in *S*-polarization is  $M_{surface,S}$  (see (7.24)). It is comprised of all the possible boundary conditions for each field. For consistency, the boundary conditions used in this matrix are

$$\hat{n} \times \vec{E} \Big|_{z \text{ at an interface}} \quad \text{and} \quad \hat{n} \times \vec{H} \Big|_{z \text{ at an interface}}$$

which can be used for both polarization cases. The matrix product of  $V_{HOMO,surface,S}$  and  $M_{surface,S}$  is the product of the inhomogeneous field boundary conditions vector  $V_{INHOMO,surface,S}$  and the inhomogeneous field amplitude,  $\mathcal{E}_{s2,S}$ , as defined by the nonlinear Fresnel expression. Explicitly, for the example of the surface region in  $S$ -polarization wave, this matrix equation is  $M_{surface,S} \cdot V_{HOMO,surface,S} = V_{INHOMO,surface,S} \mathcal{E}_{s2,S}$ . Similarly there are matrix equations for the bulk  $S$ -polarization, surface  $P$ -polarization, and bulk  $P$ -polarization waves. This section deals with the surface and bulk  $S$ -polarization waves; the following section is concerned with the  $P$ -polarization waves. In each case, the computation proceeds by explicitly calculating the elements of the homogeneous boundary condition matrix and inhomogeneous field boundary conditions vector. With the inhomogeneous field amplitude assumed known (it can be selected at a later time, when the specific numerical analyses are done), the six unknown fields shown in Figure 3.14 can be obtained. These six fields, being the elements of the homogeneous field amplitude vector, are found by solving the matrix equation. Solutions to the matrix equation were done by diagonalizing the inhomogeneous boundary condition matrix. This was done in Mathematica (for Windows, V. 2.2). The Mathematica notebook containing the solutions are reproduced in the enclosed CD-ROM; these notebooks are located in the *NLB* (nonlinear bulk) directories under the *s-wave* or *p-wave* sub-directories and are named *nlb1s.ma* and *nlb1p.ma*, respectively. The final results are reproduced in (7.43) through (7.46) for  $S$ -waves and (7.58) through (7.61) for  $P$ -waves.

### B.3.1 Determining the $S$ -polarized Electric Fields in the Layered Structure

Let the homogeneous field amplitudes (those at the generated frequency) be given by the vectors  $V_{HOMO,surface,S}$  and  $V_{HOMO,bulk,S}$  as indicated in Figure 3.14 and as described in the main text. These vectors then appear as:

$$V_{HOMO,surface,S} = \begin{pmatrix} \mathcal{E}_{r1,S} & \mathcal{E}_{t2_a,S} & \mathcal{E}_{t2_b,S} & \mathcal{E}_{u2_a,S} & \mathcal{E}_{u2_b,S} & \mathcal{E}_{v2,S} \end{pmatrix}^T, \quad (7.22)$$

$$V_{HOMO,bulk,S} = \begin{pmatrix} \mathcal{E}_{u3,S} & \mathcal{E}_{r3_a,S} & \mathcal{E}_{r3_b,S} & \mathcal{E}_{t2_a,S} & \mathcal{E}_{t3_b,S} & \mathcal{E}_{v3,S} \end{pmatrix}^T. \quad (7.23)$$

( $T$  is the transpose.) The equations that represent the satisfaction of the boundary conditions for  $S$ -waves are formed by the product of  $V_{HOMO,surface,S}$  (or  $V_{HOMO,bulk,S}$ ) with the matrices  $M_{surface,S}$  (or  $M_{bulk,S}$ ). Since the permittivity,  $\epsilon$ , and permeability,  $\mu$ , of the bulk and surface are considered scalars, the matrices have the form:

$$M_{surface,S} = \begin{pmatrix} \hat{n} \times \bar{E}_{r1}|_{z=0} & \hat{n} \times \bar{E}_{t2_a}|_{z=0} & \hat{n} \times \bar{E}_{t2_b}|_{z=0} & \hat{n} \times \bar{E}_{u2_a}|_{z=0} & \hat{n} \times \bar{E}_{u2_b}|_{z=0} & \hat{n} \times \bar{E}_{v2}|_{z=0} \\ \hat{n} \times \bar{H}_{r1}|_{z=0} & \hat{n} \times \bar{H}_{t2_a}|_{z=0} & \hat{n} \times \bar{H}_{t2_b}|_{z=0} & \hat{n} \times \bar{H}_{u2_a}|_{z=0} & \hat{n} \times \bar{H}_{u2_b}|_{z=0} & \hat{n} \times \bar{H}_{v2}|_{z=0} \\ \hat{n} \times \bar{E}_{r1}|_{z=-d_2} & \hat{n} \times \bar{E}_{t2_a}|_{z=-d_2} & \hat{n} \times \bar{E}_{t2_b}|_{z=-d_2} & \hat{n} \times \bar{E}_{u2_a}|_{z=-d_2} & \hat{n} \times \bar{E}_{u2_b}|_{z=-d_2} & \hat{n} \times \bar{E}_{v2}|_{z=-d_2} \\ \hat{n} \times \bar{H}_{r1}|_{z=-d_2} & \hat{n} \times \bar{H}_{t2_a}|_{z=-d_2} & \hat{n} \times \bar{H}_{t2_b}|_{z=-d_2} & \hat{n} \times \bar{H}_{u2_a}|_{z=-d_2} & \hat{n} \times \bar{H}_{u2_b}|_{z=-d_2} & \hat{n} \times \bar{H}_{v2}|_{z=-d_2} \\ \hat{n} \times \bar{E}_{r1}|_{z=-d_3} & \hat{n} \times \bar{E}_{t2_a}|_{z=-d_3} & \hat{n} \times \bar{E}_{t2_b}|_{z=-d_3} & \hat{n} \times \bar{E}_{u2_a}|_{z=-d_3} & \hat{n} \times \bar{E}_{u2_b}|_{z=-d_3} & \hat{n} \times \bar{E}_{v2}|_{z=-d_3} \\ \hat{n} \times \bar{H}_{r1}|_{z=-d_3} & \hat{n} \times \bar{H}_{t2_a}|_{z=-d_3} & \hat{n} \times \bar{H}_{t2_b}|_{z=-d_3} & \hat{n} \times \bar{H}_{u2_a}|_{z=-d_3} & \hat{n} \times \bar{H}_{u2_b}|_{z=-d_3} & \hat{n} \times \bar{H}_{v2}|_{z=-d_3} \end{pmatrix}_S, \quad (7.24)$$

$$M_{bulk,S} = \begin{pmatrix} \hat{n} \times \bar{E}_{u3}|_{z=0} & \hat{n} \times \bar{E}_{r3_a}|_{z=0} & \hat{n} \times \bar{E}_{r3_b}|_{z=0} & \hat{n} \times \bar{E}_{t3_a}|_{z=0} & \hat{n} \times \bar{E}_{t3_b}|_{z=0} & \hat{n} \times \bar{E}_{v3}|_{z=0} \\ \hat{n} \times \bar{H}_{u3}|_{z=0} & \hat{n} \times \bar{H}_{r3_a}|_{z=0} & \hat{n} \times \bar{H}_{r3_b}|_{z=0} & \hat{n} \times \bar{H}_{t3_a}|_{z=0} & \hat{n} \times \bar{H}_{t3_b}|_{z=0} & \hat{n} \times \bar{H}_{v3}|_{z=0} \\ \hat{n} \times \bar{E}_{u3}|_{z=-d_2} & \hat{n} \times \bar{E}_{r3_a}|_{z=-d_2} & \hat{n} \times \bar{E}_{r3_b}|_{z=-d_2} & \hat{n} \times \bar{E}_{t3_a}|_{z=-d_2} & \hat{n} \times \bar{E}_{t3_b}|_{z=-d_2} & \hat{n} \times \bar{E}_{v3}|_{z=-d_2} \\ \hat{n} \times \bar{H}_{u3}|_{z=-d_2} & \hat{n} \times \bar{H}_{r3_a}|_{z=-d_2} & \hat{n} \times \bar{H}_{r3_b}|_{z=-d_2} & \hat{n} \times \bar{H}_{t3_a}|_{z=-d_2} & \hat{n} \times \bar{H}_{t3_b}|_{z=-d_2} & \hat{n} \times \bar{H}_{v3}|_{z=-d_2} \\ \hat{n} \times \bar{E}_{u3}|_{z=-d_3} & \hat{n} \times \bar{E}_{r3_a}|_{z=-d_3} & \hat{n} \times \bar{E}_{r3_b}|_{z=-d_3} & \hat{n} \times \bar{E}_{t3_a}|_{z=-d_3} & \hat{n} \times \bar{E}_{t3_b}|_{z=-d_3} & \hat{n} \times \bar{E}_{v3}|_{z=-d_3} \\ \hat{n} \times \bar{H}_{u3}|_{z=-d_3} & \hat{n} \times \bar{H}_{r3_a}|_{z=-d_3} & \hat{n} \times \bar{H}_{r3_b}|_{z=-d_3} & \hat{n} \times \bar{H}_{t3_a}|_{z=-d_3} & \hat{n} \times \bar{H}_{t3_b}|_{z=-d_3} & \hat{n} \times \bar{H}_{v3}|_{z=-d_3} \end{pmatrix}_S. \quad (7.25)$$

For continuity of tangential  $E$  and  $H$  (recall  $\hat{n} \cdot \vec{D} = 0$  for  $S$ -polarization), these products must be equal to a vector containing the inhomogeneous fields,  $V_{INHOMO,S}$ :

$$V_{INHOMO,S} = \left( \hat{n} \times \vec{E}_{s2} \Big|_{z=0} \quad \hat{n} \times \vec{H}_{s2} \Big|_{z=0} \quad \hat{n} \times \vec{E}_{s2} \Big|_{z=d_2} \quad \hat{n} \times \vec{H}_{s2} \Big|_{z=d_2} \quad \hat{n} \times \vec{E}_{s2} \Big|_{z=d_3} \quad \hat{n} \times \vec{H}_{s2} \Big|_{z=d_3} \right)_S^T. \quad (7.26)$$

Here  $\vec{E}_{s2}$  represents the electric field induced by the nonlinear polarization and similarly for  $\vec{H}_{s2}$ .

In the expressions for the reflected fields, the following definitions will be made:

$$\sin\phi_{u2} \equiv \frac{\omega_n}{c} n_{u2} d_3 \cos\theta_{u2}, \quad (7.27)$$

$$\sin\phi_{s2} \equiv \frac{\omega_n}{c} n_{s2} d_2 \cos\theta_{s2}, \quad (7.28)$$

$$\sin\phi_{v2} \equiv \frac{\omega_n}{c} n_{v2} (d_2 + d_3) \cos\theta_{v2}, \quad (7.29)$$

$$\sin\phi_{t2} \equiv \frac{\omega_n}{c} n_{t2} d_2 \cos\theta_{t2}, \quad (7.30)$$

$$\Phi_2 \equiv \left( - \left( \bar{k}_{n,t2} - \sum_{i=1}^{n-1} \bar{k}_{i,t2} \right) \cdot \frac{\bar{k}_{n,t2}}{|\bar{k}_{n,t2}|} \frac{d_2}{2 \cos\theta_{t2}} \right) \equiv \left( - (\bar{k}_{n,t2} - \bar{k}_{n,s2}) \cdot \frac{\bar{k}_{n,t2}}{|\bar{k}_{n,t2}|} \frac{d_2}{2 \cos\theta_{t2}} \right), \quad (7.31)$$



$$\sin\phi_{u3} \equiv \frac{\omega_n}{c} n_{u3} d_2 \cos\theta_{u3}, \quad (7.32)$$

$$\sin\phi_{s3} \equiv \frac{\omega_n}{c} n_{s3} d_3 \cos\theta_{s3}, \quad (7.33)$$

$$\sin\phi_{v3} \equiv \frac{\omega_n}{c} n_{v3} (d_2 + d_3) \cos\theta_{v3}, \quad (7.34)$$

$$\sin\phi_{t3} \equiv \frac{\omega_n}{c} n_{t3} d_3 \cos\theta_{t3}, \quad (7.35)$$

and

$$\Phi_3 \equiv \left( -\left( \bar{k}_{n,t3} - \sum_{i=1}^{n-1} \bar{k}_{i,t3} \right) \cdot \frac{\bar{k}_{n,t3}}{|\bar{k}_{n,t3}|} \frac{d_3}{2 \cos\theta_{t3}} \right) \equiv \left( -\left( \bar{k}_{n,t3} - \bar{k}_{n,s3} \right) \cdot \frac{\bar{k}_{n,t3}}{|\bar{k}_{n,t3}|} \frac{d_3}{2 \cos\theta_{t3}} \right). \quad (7.36)$$

Realizing that  $\bar{E}_{r1} = \bar{e}_{r1} \mathcal{E}_{r1} e^{i\bar{k}_{r1} \cdot \bar{r}}$  for harmonic field  $\bar{E}_{r1}(\omega_n)$  and similarly for all the other fields, the matrices can be written:

$$M_{surface,S} = \begin{pmatrix} -1 & 1 & 1 & 0 & 0 & 0 \\ +n_{r1} \cos\theta_{r1} & +n_{r2} \cos\theta_{r2} & -n_{r2} \cos\theta_{r2} & 0 & 0 & 0 \\ 0 & -e^{i\theta_{r2}} & -e^{-i\theta_{r2}} & 1 & 1 & 0 \\ 0 & -n_{r2} \cos\theta_{r2} e^{i\theta_{r2}} & +n_{r2} \cos\theta_{r2} e^{-i\theta_{r2}} & +n_{u2} \cos\theta_{u2} & -n_{u2} \cos\theta_{u2} & 0 \\ 0 & 0 & 0 & -e^{i\theta_{u2}} & -e^{-i\theta_{u2}} & 1 \\ 0 & 0 & 0 & -n_{u2} \cos\theta_{u2} e^{i\theta_{u2}} & +n_{u2} \cos\theta_{u2} e^{-i\theta_{u2}} & +n_{v2} \cos\theta_{v2} \end{pmatrix}, \quad (7.37)$$

$$M_{bulk,S} = \begin{pmatrix} -1 & 1 & 1 & 0 & 0 & 0 \\ +n_{s3}\cos\theta_{s3} & -n_{r3}\cos\theta_{r3}e^{i\phi_{r3}} & +n_{r3}\cos\theta_{r3}e^{-i\phi_{r3}} & 0 & 0 & 0 \\ 0 & -1 & -1 & 1 & 1 & 0 \\ 0 & +n_{r3}\cos\theta_{r3} & -n_{r3}\cos\theta_{r3} & +n_{t3}\cos\theta_{t3} & -n_{t3}\cos\theta_{t3} & 0 \\ 0 & 0 & 0 & -e^{i\phi_{t3}} & -e^{-i\phi_{t3}} & 1 \\ 0 & 0 & 0 & -n_{t3}\cos\theta_{t3}e^{i\phi_{t3}} & +n_{t3}\cos\theta_{t3}e^{-i\phi_{t3}} & +n_{v3}\cos\theta_{v3} \end{pmatrix}, \quad (7.38)$$

and the inhomogeneous field vectors are found to be:

$$V_{INHOMO,surface,S} = \begin{pmatrix} 1 & n_{s2}\cos\theta_{s2} & -e^{i\phi_{s2}} & -n_{s2}\cos\theta_{s2}e^{i\phi_{s2}} & 0 & 0 \end{pmatrix}^T, \quad (7.39)$$

$$V_{INHOMO,bulk,S} = \begin{pmatrix} 0 & 0 & 1 & n_{t3}\cos\theta_{t3} & -e^{i\phi_{t3}} & -n_{t3}\cos\theta_{t3}e^{i\phi_{t3}} \end{pmatrix}^T. \quad (7.40)$$

Thus, to solve for the reflected and transmitted fields, solutions to the matrix equations

$$M_{surface,S} \cdot V_{HOMO,surface,S} = V_{INHOMO,surface,S} \mathcal{E}_{s2,S} \quad (7.41)$$

and

$$M_{bulk,S} \cdot V_{HOMO,bulk,S} = V_{INHOMO,bulk,S} \mathcal{E}_{s3,S} \quad (7.42)$$

must be found. Solutions to these equations are found by diagonalizing the composite matrices. The solutions for the  $S$ -polarization surface reflected field,  $\mathcal{E}_{r1,S}$ , and bulk

reflected field,  $\mathcal{E}_{u3,S}$ , are found on the following pages;

$$\mathcal{E}_{r1,S} = D_{r1,S}^{-1} \mathcal{E}_{s2,S} \left\{ \begin{array}{l} n_{u2} \cos \theta_{u2} \cos \phi_{u2} \left( \begin{array}{l} n_{s2} \cos \theta_{s2} \\ -n_{v2} \cos \theta_{v2} \end{array} \right) (\cos \phi_{s2} - \cos \phi_{i2}) \\ + n_{u2}^2 \cos^2 \theta_{u2} \sin \phi_{u2} \left( \sin \phi_{s2} - \frac{n_{s2}}{n_{i2}} \cos \theta_{s2} \sec \theta_{i2} \sin \phi_{i2} \right) \\ + n_{v2} \cos \theta_{v2} \sin \phi_{v2} (n_{s2} \cos \theta_{s2} \sin \phi_{s2} - n_{i2} \cos \theta_{i2} \sin \phi_{i2}) \\ + i n_{u2} \cos \theta_{u2} \cos \phi_{u2} \sin \phi_{s2} \left( \begin{array}{l} n_{s2} \cos \theta_{s2} \\ -n_{v2} \cos \theta_{v2} \end{array} \right) \\ - i n_{u2} \cos \theta_{u2} \cos \phi_{u2} \sin \phi_{i2} \\ \times \left( \begin{array}{l} n_{i2} \cos \theta_{i2} - \\ \frac{n_{s2} n_{v2}}{n_{i2}} \cos \theta_{s2} \cos \theta_{v2} \sec \theta_{i2} \end{array} \right) \\ - i n_{u2} \cos \theta_{u2} \sin \phi_{u2} (\cos \phi_{s2} - \cos \phi_{i2}) \\ \times \left( n_{v2} \cos \theta_{v2} - \frac{n_{s2} n_{v2}}{n_{i2}} \cos \theta_{s2} \cos \theta_{v2} \sec \theta_{u2} \right) \end{array} \right\} \quad (7.43)$$

with

$$D_{r1,S} = \left\{ \begin{array}{l} n_{u2} \cos \theta_{u2} \cos \phi_{u2} (n_{r1} \cos \theta_{r1} \cos \phi_{i2} + n_{v2} \cos \theta_{v2} \cos \phi_{i2}) \\ - \left( \begin{array}{l} n_{i2} n_{v2} \cos \theta_{i2} \cos \theta_{v2} \\ - \frac{n_{r1} n_{u2}^2}{n_{i2}} \cos \theta_{r1} \cos^2 \theta_{u2} \sec \theta_{i2} \end{array} \right) \\ - i n_{u2} \cos \theta_{u2} \left( \begin{array}{l} \cos \theta_{i2} \cos \phi_{u2} \sin \phi_{i2} \\ + n_{u2} \cos \theta_{u2} \sin \phi_{u2} \cos \phi_{i2} \end{array} \right) \\ - i n_{r1} n_{v2} \cos \theta_{r1} \cos \theta_{v2} \left( \begin{array}{l} \frac{n_{u2}}{n_{i2}} \cos \theta_{u2} \sec \theta_{i2} \cos \phi_{u2} \sin \phi_{i2} \\ - \cos \phi_{i2} \sin \phi_{u2} \end{array} \right) \end{array} \right\} \quad (7.44)$$

$$\mathcal{E}_{u3,S} = D_{u3,S}^{-1} \mathcal{E}_{s3,S} \left\{ \begin{array}{l} \frac{n_{t3}}{n_{r3}} \cos \theta_{t3} \sec \theta_{r3} \cos \phi_{r3} \begin{pmatrix} n_{s3} \cos \theta_{s3} \\ -n_{v3} \cos \theta_{v3} \end{pmatrix} (\cos \phi_{s3} - \cos \phi_{t3}) \\ + i \frac{n_{t3}}{n_{r3}} \cos \theta_{t3} \sec \theta_{r3} \cos \phi_{r3} \begin{pmatrix} n_{s3} \cos \theta_{s3} \sin \phi_{s3} \\ -n_{v3} \cos \theta_{v3} \sin \phi_{s3} \\ -n_{t3} \cos \theta_{t3} \sin \phi_{t3} \\ + \frac{n_{s3} n_{v3}}{n_{t3}} \cos \theta_{s3} \cos \theta_{v3} \\ \times \sec \theta_{t3} \sin \phi_{t3} \end{pmatrix} \end{array} \right\} \quad (7.45)$$

with

$$D_{u3,S} = \left\{ \begin{array}{l} n_{t3} \cos \theta_{t3} \begin{pmatrix} n_{s3} \cos \theta_{s3} \\ -n_{v3} \cos \theta_{v3} \end{pmatrix} \begin{pmatrix} \cos \phi_{t3} \sin \phi_{r3} \\ -\frac{n_{t3}}{n_{r3}} \cos \theta_{t3} \sin \phi_{t3} \sec \theta_{r3} \cos \phi_{r3} \end{pmatrix} \\ -i \frac{n_{u3} n_{v3}}{n_{r3}} \cos \theta_{u3} \cos \theta_{v3} \sec \theta_{r3} \sin \phi_{t3} \end{array} \right\} \quad (7.46)$$

#### B.4 Determining the $P$ -polarized Electric Fields in the Layered Structure

Let the homogeneous field amplitudes (those at the generated frequency) be given by the vectors  $V_{HOMO,surface,P}$  and  $V_{HOMO,bulk,P}$  as indicated in Figure 3.14. These vectors then appear as:

$$V_{HOMO,surface,P} = \begin{pmatrix} \mathcal{E}_{r1,P} & \mathcal{E}_{t2a,P} & \mathcal{E}_{t2b,P} & \mathcal{E}_{u2a,P} & \mathcal{E}_{u2b,P} & \mathcal{E}_{v2,P} \end{pmatrix}^T, \quad (7.47)$$

$$V_{HOMO,bulk,P} = \begin{pmatrix} \mathcal{E}_{u3,P} & \mathcal{E}_{r3a,P} & \mathcal{E}_{r3b,P} & \mathcal{E}_{t2a,P} & \mathcal{E}_{t3b,P} & \mathcal{E}_{v3,P} \end{pmatrix}^T. \quad (7.48)$$

Recalling that for  $P$ -waves  $\hat{n} \cdot \vec{B} = 0$ . The equations that satisfy the boundary conditions for  $P$ -waves are formed by the product of  $V_{HOMO,surface,P}$  (or  $V_{HOMO,bulk,P}$ ) with the matrices  $M_{surface,P}$  (or  $M_{bulk,P}$ ). These have the form:

$$M_{surface,P} = \begin{pmatrix} \hat{n} \times \vec{E}_{r1}|_{z=0} & \hat{n} \times \vec{E}_{t2a}|_{z=0} & \hat{n} \times \vec{E}_{t2b}|_{z=0} & \hat{n} \times \vec{E}_{u2a}|_{z=0} & \hat{n} \times \vec{E}_{t2b}|_{z=0} & \hat{n} \times \vec{E}_{v2}|_{z=0} \\ \hat{n} \times \vec{H}_{r1}|_{z=0} & \hat{n} \times \vec{H}_{t2a}|_{z=0} & \hat{n} \times \vec{H}_{t2b}|_{z=0} & \hat{n} \times \vec{H}_{u2a}|_{z=0} & \hat{n} \times \vec{H}_{t2b}|_{z=0} & \hat{n} \times \vec{H}_{v2}|_{z=0} \\ \hat{n} \times \vec{E}_{r1}|_{z=-d_2} & \hat{n} \times \vec{E}_{t2a}|_{z=-d_2} & \hat{n} \times \vec{E}_{t2b}|_{z=-d_2} & \hat{n} \times \vec{E}_{u2a}|_{z=-d_2} & \hat{n} \times \vec{E}_{t2b}|_{z=-d_2} & \hat{n} \times \vec{E}_{v2}|_{z=-d_2} \\ \hat{n} \times \vec{H}_{r1}|_{z=-d_2} & \hat{n} \times \vec{H}_{t2a}|_{z=-d_2} & \hat{n} \times \vec{H}_{t2b}|_{z=-d_2} & \hat{n} \times \vec{H}_{u2a}|_{z=-d_2} & \hat{n} \times \vec{H}_{t2b}|_{z=-d_2} & \hat{n} \times \vec{H}_{v2}|_{z=-d_2} \\ \hat{n} \times \vec{E}_{r1}|_{z=-d_3} & \hat{n} \times \vec{E}_{t2a}|_{z=-d_3} & \hat{n} \times \vec{E}_{t2b}|_{z=-d_3} & \hat{n} \times \vec{E}_{u2a}|_{z=-d_3} & \hat{n} \times \vec{E}_{t2b}|_{z=-d_3} & \hat{n} \times \vec{E}_{v2}|_{z=-d_3} \\ \hat{n} \times \vec{H}_{r1}|_{z=-d_3} & \hat{n} \times \vec{H}_{t2a}|_{z=-d_3} & \hat{n} \times \vec{H}_{t2b}|_{z=-d_3} & \hat{n} \times \vec{H}_{u2a}|_{z=-d_3} & \hat{n} \times \vec{H}_{t2b}|_{z=-d_3} & \hat{n} \times \vec{H}_{v2}|_{z=-d_3} \end{pmatrix}_P, \quad (7.49)$$

$$M_{bulk,P} = \begin{pmatrix} \hat{n} \times \vec{E}_{u3}|_{z=0} & \hat{n} \times \vec{E}_{r3a}|_{z=0} & \hat{n} \times \vec{E}_{r3b}|_{z=0} & \hat{n} \times \vec{E}_{t3a}|_{z=0} & \hat{n} \times \vec{E}_{t3b}|_{z=0} & \hat{n} \times \vec{E}_{v3}|_{z=0} \\ \hat{n} \times \vec{H}_{u3}|_{z=0} & \hat{n} \times \vec{H}_{r3a}|_{z=0} & \hat{n} \times \vec{H}_{r3b}|_{z=0} & \hat{n} \times \vec{H}_{t3a}|_{z=0} & \hat{n} \times \vec{H}_{t3b}|_{z=0} & \hat{n} \times \vec{H}_{v3}|_{z=0} \\ \hat{n} \times \vec{E}_{u3}|_{z=-d_2} & \hat{n} \times \vec{E}_{r3a}|_{z=-d_2} & \hat{n} \times \vec{E}_{r3b}|_{z=-d_2} & \hat{n} \times \vec{E}_{t3a}|_{z=-d_2} & \hat{n} \times \vec{E}_{t3b}|_{z=-d_2} & \hat{n} \times \vec{E}_{v3}|_{z=-d_2} \\ \hat{n} \times \vec{H}_{u3}|_{z=-d_2} & \hat{n} \times \vec{H}_{r3a}|_{z=-d_2} & \hat{n} \times \vec{H}_{r3b}|_{z=-d_2} & \hat{n} \times \vec{H}_{t3a}|_{z=-d_2} & \hat{n} \times \vec{H}_{t3b}|_{z=-d_2} & \hat{n} \times \vec{H}_{v3}|_{z=-d_2} \\ \hat{n} \times \vec{E}_{u3}|_{z=-d_3} & \hat{n} \times \vec{E}_{r3a}|_{z=-d_3} & \hat{n} \times \vec{E}_{r3b}|_{z=-d_3} & \hat{n} \times \vec{E}_{t3a}|_{z=-d_3} & \hat{n} \times \vec{E}_{t3b}|_{z=-d_3} & \hat{n} \times \vec{E}_{v3}|_{z=-d_3} \\ \hat{n} \times \vec{H}_{u3}|_{z=-d_3} & \hat{n} \times \vec{H}_{r3a}|_{z=-d_3} & \hat{n} \times \vec{H}_{r3b}|_{z=-d_3} & \hat{n} \times \vec{H}_{t3a}|_{z=-d_3} & \hat{n} \times \vec{H}_{t3b}|_{z=-d_3} & \hat{n} \times \vec{H}_{v3}|_{z=-d_3} \end{pmatrix}_P. \quad (7.50)$$

These products must be equal to a vector containing the inhomogeneous fields,  $V_{INHOMO,S}$ . This vector appears as

$$V_{INHOMO,P} = \left( \hat{n} \times \bar{E}_{s2}|_{z=0} \quad \hat{n} \times \bar{H}_{s2}|_{z=0} \quad \hat{n} \times \bar{E}_{s2}|_{z=-d_2} \quad \hat{n} \times \bar{H}_{s2}|_{z=-d_2} \quad \hat{n} \times \bar{E}_{s2}|_{z=-d_3} \quad \hat{n} \times \bar{H}_{s2}|_{z=-d_3} \right)_P^T. \quad (7.51)$$

Here  $\bar{E}_{s2}$  represents the electric field induced by the nonlinear polarization and similarly for  $\bar{H}_{s2}$ .

Realizing that  $\bar{E}_{r1} = \bar{E}_{r1} \mathcal{E}_{r1} e^{i\vec{k}_{r1} \cdot \vec{r}}$  for harmonic field  $\bar{E}_{r1}(\omega_n)$  and similarly for all the other fields, the matrices can be written:

$$M_{surface,P} = \begin{pmatrix} \cos\theta_{r1} & \cos\theta_{r2} & -\cos\theta_{r2} & 0 & 0 & 0 \\ -n_{r1} & n_{r2} & n_{r2} & 0 & 0 & 0 \\ 0 & -\cos\theta_{r2}e^{i\theta_{r2}} & \cos\theta_{r2}e^{-i\theta_{r2}} & \cos\theta_{s2} & -\cos\theta_{s2} & 0 \\ 0 & -n_{r2}e^{i\theta_{r2}} & -n_{r2}e^{-i\theta_{r2}} & n_{s2} & n_{s2} & 0 \\ 0 & 0 & 0 & -\cos\theta_{s2}e^{i\theta_{s2}} & \cos\theta_{s2}e^{-i\theta_{s2}} & \cos\theta_{v2} \\ 0 & 0 & 0 & -n_{s2}e^{i\theta_{s2}} & -n_{s2}e^{-i\theta_{s2}} & n_{v2} \end{pmatrix}, \quad (7.52)$$

$$M_{bulk,P} = \begin{pmatrix} \cos\theta_{s3} & \cos\theta_{r3}e^{i\theta_{r3}} & -\cos\theta_{r3}e^{-i\theta_{r3}} & 0 & 0 & 0 \\ -n_{s3} & n_{r3}e^{i\theta_{r3}} & n_{r3}e^{-i\theta_{r3}} & 0 & 0 & 0 \\ 0 & \cos\theta_{r3} & -\cos\theta_{r3} & \cos\theta_{i3} & -\cos\theta_{i3} & 0 \\ 0 & -n_{r3} & -n_{r3} & n_{i3} & n_{i3} & 0 \\ 0 & 0 & 0 & -\cos\theta_{i3}e^{i\theta_{i3}} & \cos\theta_{i3}e^{-i\theta_{i3}} & \cos\theta_{v3} \\ 0 & 0 & 0 & -n_{i3}e^{i\theta_{i3}} & -n_{i3}e^{-i\theta_{i3}} & n_{v3} \end{pmatrix}, \quad (7.53)$$

and the inhomogeneous field vectors are found to be:

$$V_{INHOMO,surface,P} = \left( \mathcal{E}_D \quad \mathcal{E}_B \quad -\mathcal{E}_D e^{i\theta_{r2}} \quad -\mathcal{E}_B e^{i\theta_{r2}} \quad 0 \quad 0 \right)^T, \quad (7.54)$$

$$V_{INHOMO,bulk,P} = \left( 0 \quad 0 \quad \mathcal{E}_D \quad \mathcal{E}_B \quad -\mathcal{E}_D e^{i\theta_{r3}} \quad -\mathcal{E}_B e^{i\theta_{r3}} \right)^T. \quad (7.55)$$

Thus, to solve for the reflected and transmitted fields, solutions to the matrix equations

$$M_{surface,P} \cdot V_{HOMO,surface,P} = V_{INHOMO,surface,P} \mathcal{E}_{s2,P} \quad (7.56)$$

and

$$M_{bulk,P} \cdot V_{HOMO,bulk,P} = V_{INHOMO,bulk,P} \mathcal{E}_{s3,P} \quad (7.57)$$

must be found. Solutions to these equations are found by diagonalizing the composite matrices. The solutions are given on the next pages;

$$\mathcal{E}_{r1,P} = D_{r1,P}^{-1} \mathcal{E}_{s2,P} \left\{ \begin{array}{l} n_{u2} \cos \theta_{u2} \cos \phi_{u2} \left[ (\mathcal{E}_{B2} - \mathcal{E}_{D2}) \sin \phi_{s2} - \left( \mathcal{E}_{B2} \frac{\cos \theta_{t2}}{n_{t2}} - \mathcal{E}_{D2} \sin \phi_{s2} \right) \sin \phi_{t2} \right] \\ - \cos^2 \theta_{u2} \sin \phi_{u2} [\mathcal{E}_{B2} (\sin \phi_{s2} - \cos \phi_{t2})] \\ + n_{u2}^2 \sin \phi_{u2} [\mathcal{E}_{D2} (\cos \phi_{s2} - \cos \phi_{t2})] \\ - i n_{u2} \cos \theta_{u2} \cos \phi_{u2} [(\mathcal{E}_{B2} - \mathcal{E}_{D2}) (\cos \phi_{s2} - \cos \phi_{t2})] \\ - i \cos^2 \theta_{u2} \sin \phi_{u2} [\mathcal{E}_{B2} \sin \phi_{s2} - \mathcal{E}_{D2} \sec \theta_{t2} \sin \phi_{t2}] \\ + i n_{u2}^2 \sin \phi_{u2} \left[ \mathcal{E}_{D2} \sin \phi_{s2} - \mathcal{E}_{B2} \frac{\cos \theta_{t2}}{n_{t2}} \sin \phi_{t2} \right] \end{array} \right\} \quad (7.58)$$

with

$$D_{r1,P} = \left\{ \begin{aligned} & n_{u2} \cos \theta_{r1} (n_{i2} \cos \theta_{u2} \cos \phi_{u2} \sec \theta_{i2} \sin \phi_{i2} + n_{u2} \sin \phi_{u2} \cos \phi_{i2}) \\ & + \frac{n_{r1} n_{u2}}{n_{i2}} \cos \theta_{i2} \cos \theta_{u2} \cos \phi_{u2} \sin \phi_{i2} \\ & + n_{r1} \cos^2 \theta_{u2} \sin \phi_{u2} \cos \phi_{i2} \\ & + i n_{u2} \cos \theta_{u2} (n_{r1} + \cos \theta_{r1}) \cos \phi_{u2} \cos \phi_{i2} \\ & - i \frac{n_{r1} n_{u2}^2}{n_{i2}} \cos \theta_{i2} \sin \phi_{u2} \sin \phi_{i2} \\ & - i n_{i2} \cos \theta_{r1} \cos^2 \theta_{u2} \sec \theta_{i2} \sin \phi_{u2} \sin \phi_{i2} \end{aligned} \right\} \quad (7.59)$$

$$\mathcal{E}_{u3,P} = D_{u3,P}^{-1} \mathcal{E}_{s3,P} \left\{ \begin{aligned} & \frac{\mathcal{E}_{B3}}{4} [\cos(2\theta_{i3} - \phi_{i3}) + \cos(2\theta_{i3} + \phi_{i3}) - 2\cos \phi_{i3}] \\ & + \frac{\mathcal{E}_{B3}}{2} [\cos(\theta_{i3} - \phi_{i3}) + \cos(\theta_{i3} + \phi_{i3})] \\ & + \frac{n_{i3}}{2} (\mathcal{E}_{B3} - \mathcal{E}_{D3}) [\cos(\theta_{i3} - \phi_{i3}) - \cos(\theta_{i3} + \phi_{i3})] \\ & - \frac{i}{2} (\mathcal{E}_{B3} - \mathcal{E}_{D3}) [\sin(\theta_{i3} - \phi_{i3}) - \sin(\theta_{i3} + \phi_{i3})] \\ & - \frac{i}{2} \mathcal{E}_{D3} [\sin(\theta_{i3} - \phi_{i3}) - \sin(\theta_{i3} + \phi_{i3}) + 2n_{i3}^2 \sin \phi_{i3}] \end{aligned} \right\} \quad (7.60)$$

with

$$D_{u3,P} = \left\{ \begin{aligned} & \left( \begin{aligned} & n_{u3} \cos \theta_{i3} \cos \phi_{r3} \cos \phi_{i3} \\ & - \frac{n_{i3} n_{u3}}{n_{r3}} \cos \theta_{r3} \sin \phi_{r3} \sin \phi_{i3} \end{aligned} \right) (\cos \theta_{i3} - n_{i3}) \\ & - i \left( \begin{aligned} & n_{i3} \cos \theta_{u3} \cos \phi_{r3} \sin \phi_{i3} \\ & - n_{r3} \cos \theta_{u3} \cos \theta_{i3} \cos \phi_{i3} \sec \theta_{r3} \sin \phi_{r3} \end{aligned} \right) (\cos \theta_{i3} - n_{i3}) \end{aligned} \right\}, \quad (7.61)$$



$$\begin{aligned}
\mathcal{E}_{D2} &= \cos\theta_{s2} \sin\alpha, \\
\mathcal{E}_{B2} &= \left( \frac{n_{t2}^2 - n_{s2}^2}{n_{t2}^2} \right) \sin\theta_{s2} \cos\alpha, \\
\mathcal{E}_{D3} &= \cos\theta_{s3} \sin\alpha, \text{ and} \\
\mathcal{E}_{B3} &= \left( \frac{n_{t3}^2 - n_{s3}^2}{n_{t3}^2} \right) \sin\theta_{s3} \cos\alpha.
\end{aligned} \tag{7.62}$$

### B.5 Connections to Bloembergen's Results

It is important to note that if the bulk median layer is removed from this model structure, it is reduced to Bloembergen's model. Hence, if this layer is removed from these derivations, the expressions found here should reduce to those quoted by Bloembergen. To explicitly remove the bulk median layer, the rows in  $V_{HOMO,surface,S}$  which contain  $\mathcal{E}_{u2_a,S}$  and  $\mathcal{E}_{u2_b,S}$  are eliminated. Also  $V_{HOMO,bulk,S}$  is set to zero (there is no longer a bulk median layer for a generated signal to arise in). This argument was made explicitly for  $S$ -polarizations, but can also be made for  $P$ -polarizations. The solutions to the resulting matrix equations (both for  $S$  and  $P$ ) indeed reproduce Bloembergen's work. To make direct comparisons with Bloembergen's notation, the following subscripts change:  $r1 \rightarrow r$ ,  $t2 \rightarrow m$ ,  $s2 \rightarrow s$ , and  $v3 \rightarrow t$ . Also the substitution,  $\phi_2 \rightarrow n_m/c d \cos \theta_m$  must be made.

Likewise, looking at the solutions to the reflected (surface) field amplitude,  $\mathcal{E}_{r1,S}$ , Bloembergen's result is returned if  $n_3 \rightarrow n_4 \rightarrow n_m$ ,  $n_{s3} \rightarrow 0$ ,  $n_4 \rightarrow n_t$  and  $\phi_2 \rightarrow n_m/c d \cos \theta_m \rightarrow \phi_m$ . These substitutions have the affect of removing the bulk median layer by letting the index of refraction revert to that of the semi-infinite bulk and removing the phase-(mis)matching information. Equivalently, setting  $d_3 = d_2$  returns Bloembergen's

expression for the reflected field. Each of these tests of removing the bulk median layer and getting back Bloembergen's expressions lends credence to the expressions derived here.

## B.6 Summary of the Field Expressions for the Macroscopic RISOM Model

The solutions presented here are quite general. Bloembergen, in his model, approximated the thickness of the surface region as small compared with the generated wavelength. This approximation is not done here. Moreover, the exact phase-matching expression is retained. The solutions to the field expressions for this macroscopic RISOM model were found without exact knowledge of the induced nonlinear polarization. Hence, the solutions obtained are independent of the OM process (the order of the nonlinear polarization) and are general with respect to the incident wave parameters. Hence, the reflection coefficients, (7.43) through (7.46) and (7.58) through (7.61), for surface and bulk were derived for *general optical mixing*. This was done by defining a generalized effective source wavevector and source index of refraction (7.1) which are due to the incident fields (and linear properties of the media). From the equations for general mixing, a simple substitution can be made to predict RISOM for specific processes such as CARS-like FWM (four-wave mixing), SHG (second harmonic generation), and TWM (three-wave mixing).

An important separation of the various dependencies on angle and phase-mismatch can be made. For instance, consider an *S*-wave generated in the surface layer. The amplitude of this field observed in the vacuum has the form

$$\begin{aligned}\mathcal{E}_{observed,S} &= \mathcal{E}_{r1,S} = \mathcal{E}_{BC2,S} \mathcal{E}_{s,S}(\theta_{i2}, \Phi_2) \\ &= \mathcal{E}_{BC2,S} \left( \frac{-4\pi |\bar{P}_S|}{n_t^2 - n_s^2} \right) \sec \theta_{i2} \operatorname{sinc} \Phi_2 e^{-i\Phi},\end{aligned}$$

where  $\mathcal{E}_{BC2,S} = \mathcal{E}_{r1,S} / \mathcal{E}_{s2,S}$  (see (7.43)). This form clearly shows that all the polarization dependence is in  $\mathcal{E}_{BC2,S}$  (recalling  $|\bar{P}_S|$  is taken to be equal to  $|\bar{P}_S|$ ). The  $\sec \theta_{i2}$  dependence is in addition to the angular dependence in  $\mathcal{E}_{BC2,S}$ . The dependence of the observed field strength on mismatch shows up in  $\mathcal{E}_{BC2,S}$ , is modified by  $\text{sinc } \Phi_2$ , but shows its strongest dependence on  $(n_i^2 - n_s^2)^{-1}$ .