## INFORMATION TO USERS

While the most advanced technology has been used to photograph and reproduce this manuscript, the quality of the reproduction is heavily dependent upon the quality of the material submitted. For example:

- Manuscript pages may have indistinct print. In such cases, the best available copy has been filmed.
- Manuscripts may not always be complete. In such cases, a note will indicate that it is not possible to obtain missing pages.
- Copyrighted material may have been removed from the manuscript. In such cases, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, and charts) are photographed by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps. Each oversize page is also filmed as one exposure and is available, for an additional charge, as a standard 35 mm slide or as a 17 " 23 " black and white photographic print.

Most photographs reproduce acceptably on positive microfilm or microfiche but lack the clarity on xerographic copies made from the microfilm. For an additional charge, 35 mm slides of 6 "x 9 " black and white photographic prints are available for any photographs or illustrations that cannot be reproduced satisfactorily by xerography.

$$
r
$$

# Chipman, Russell Atwood 

POLARIZATION ABERRATIONS

The University of Arizona
Рн.D. 1987
University
Microfilms
International 300 N . Zees Road. Am Ambor, M48106
Copyright 1987
by
Chipman, Russell Atwood
All Rights Reserved
$!$

## PLEASE NOTE:

In all cases this material has been filmed in the best possible way from the available copy. Problems encountered with this document have been identified here with a check mark $\qquad$ 1.

1. Glossy photographs or pages $\qquad$
2. Colored illustrations, paper or print $\qquad$
3. Photographs with dark background $\qquad$
4. Illustrations are poor copy $\qquad$
5. Pages with black marks, not original copy $\qquad$
6. Print shows through as there is text on both sides of page $\qquad$
7. Indistinct, broken or small print on several pages $\qquad$
8. Print exceeds margin requirements $\qquad$
9. Tightly bound copy with print lost in spine $\qquad$
10. Computer printout pages with indistinct print $\qquad$
11. Page(s) $\qquad$ lacking when material received, and not available from school or author.
12. Page(s) $\qquad$ seem to be missing in numbering only as text foilows.
13. Two pages numbered $\qquad$ . Text follows.
14. Curling and wrinkled pages $\qquad$
15. Dissertation contains pages with print at a slant, filmed as received $\qquad$
16. Other $\qquad$
$\qquad$

- 


# POLARIZATION ABERRATIONS 

by
Russell Atwood Chipman

A Dissertation Submitied to the Faculty of the COMMITTEE ON OPTICAL SCIENCES (GRADUATE)

In Partial Fulfillment of the Requirements
For the Degree of
DOCTOR OF PHILOSOPHY
In the Graduate College
THE UNIVERSITY OF ARIZONA

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Russell Atwood Chipman

```
entitled Polarization Aberrations
```

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy -

Date

Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that $I$ have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

$4 / 21 / 87$
Date

## STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of source is made. Requests for permission for extended quotation from or reproduction of this manuscript in whole or in part may be granted by the copyright holder.


## PREFACE

This dissertation is dedicated to my wife Laure for her constant support and enthusiasm.

The work was supported at different stages by two NASA laboratories. The Jet Propulsion Laboratory initially supported the writing a polarization ray tracing program to calculate instrumental polarization. This ray tracing work eventually led to the present concepis of polarization aberrations. In particular, Dr. James Breckenridge at JPL suggested the area of instrumental poiarization as fertile for research and he provided substantial suppori. Dr. John Stacy at JPL also provided much assistance. Later, the research was supported by Dr. Mona Hagyard and Dr. Alan Gary at the Marshall Space Flight Center who offered a particularly challenging optical system in need of polarization aberration analysis, a Solar Vector Magnetograph for the determination of solar magnetic fields through the measurement of Zeeman splitting.

Several of my professors at the Optical Sciences Center made noteworthy contributions to this research: Dr. James Wyant for supervising this project and teaching me so much about optics, Dr. Roland Shack for his simple and straightforward explanations of aberration theory. Dr. Oersted Stavroudis for his unique understanding of optical propagation. Dr. Angus Macleod for his explanations of thin films and Robert Shannon for his knowledge and good sense about optical design. I also wish to thank my colleagues Dr. Jean Bennett and James P. McGuire Jr. at the Center for Applied Optics at the University of Alabama in Huntsville for
many suggestions about the material and for their careful reading of the manuscript. Finally, I wish to acknowledge Dr. R. Clark Jones for his solid foundation of research into polarizers upon which this theoretical treatment of polarization aberrations is constructed.

## TABLE OF CONTENTS

Page
LIST OF TABLES ..... X
LIST OF ILLUSTRATIONS ..... $x i$
ABSTRACT ..... xii

1. INIRODUCTION ..... 1
Polarized Light ..... 2
Polarization Elements ..... 2
The Eigit Forms of Poiarization Benavior ..... 3
Instrumental and Residual Polarization ..... 3
Sources of Instrumental Polarization ..... 5
Systems Sensitive to Instrumental Polarization ..... 7
Polarization Aberrations ..... 9
Optical Design ..... 12
Thin Films ..... 17
The Wavefront Aberration Function ..... 19
The Instrumental Polarization Function ..... 20
The Jones Calculus Representation of the
Instrumental Polarization Function ..... 21
The Jones Matrix Form for Interface Polarization ..... 22
The Jones Matrix for Propagation ..... 24
The Jones Matrix for Ray Paths Through Optical Systems ..... 24
Interpretation of Polarization Aberration Results ..... 25
2. THE JONES MATRIX AND C VECTOR FOR THE CHARACTERIZATION OF POLARIZATION ..... 27
The Jones Calculus ..... 27
Polarization Elements ..... 28
Notation ..... 29
The Jones Vector ..... 29
The Coherent Addition of Optical Fields ..... 30
Relation to Elliptical Polarization Parameters ..... 31
Basis Jones Vectors ..... 32
Generalizations of the Jones Vector ..... 32
The Jones Matrix and the C Vector ..... 34
Definition of the Jones Matrix in terms of the Jones Vecior ..... 34
Cascaded Polarizers ..... 34
Coordinate System ..... 35
Basis States ..... 36

## TABLE OF CONTENTS--Continued

Page
Pauli Spin Matrix Basis and the C Vector ..... 37
The Jones Matrix and C Vectors for Specific Polarizers ..... 39
Normalized Form for the C Vectors ..... 40
Pauli Spin Matrices ..... 40
The Identity Matrix ..... 40
Trace and Determinant ..... 44
Anticommutativity ..... 44
Relation to the Identity Matrix ..... 44
Eigenvalues and Eigenvectors ..... 44
Effect of the Pauli Spin Matrices on Basis Polarization States ..... 46
The Exponential of a Matrix ..... 46
Exponentials of the Basis Matrices ..... 48
Properties of the Jones Matrix and the C Vector ..... 48
Rotated Polarizers ..... 48
Rotations of the Basis Matrices ..... 49
Rotation of a C Vector ..... 49
The C Vector for Cascaded Polarizers ..... 50
The Meaning of the Coefficients of the $C$ Vector ..... 51
3. WEAK POLARIZERS ..... 53
Instrumental Polarization versus Transmitted Light Polarization Calculations ..... 53
Transparent Systems ..... 54
S-P Coordinates ..... 54
PIormalized C Vector ..... 56
Instrumental Polarization ..... 57
Weak Polarizers ..... 57
Amplitude Transmission Relations ..... 58
Taylor Series Representation for Weak Polarizers ..... 60
Obtaining Taylor Series Coefficients from Sampled Data ..... 62
The Fresnel Equations ..... 63
Intensity and Power Transmission Relations ..... 68
Refraction, The Transmitted Beam ..... 69
Single Layer Dielectric Thin Films ..... 71
Antireflection Coating Example ..... 75
4. CASCADED WEAK POLARIZERS ..... 83
Introduction ..... 83
The Order Dependence of Pairs of Polarizers and Retarders ..... 84
Pairs of Linear Retarders ..... 84
Example ..... 85

## TABLE OF CONTENTS--Continued

Page
Example ..... 86
Pairs of Linear Polarizers ..... 86
Pairs of Linear Polarizers and Linear Retarders ..... 87
Pairs Involving Circular Polarizers and Circular Retarders ..... 88
Summary of Polarization Couplings ..... 88
Instrumental Polarization for Ray Paths through Optical Systems ..... 88
General Case ..... 89
Homogeneous Optical System ..... 89
Radially Symmetric Systems of Lenses, Mirrors and Coatings ..... 91
The C Vector for a Paraxial Ray ..... 93
5. RESULTS FROM GEOMETRICAL OPTICS ..... 96
Introduction ..... 96
Paraxial Optics Summary ..... 96
Coordinate System ..... 97
Subscripts ..... 97
The Paraxial Angle of Incidence ..... 99
The Wavefront Aberrations ..... 99
The Lowest Order Terms ..... 102
Constant Piston Error ..... 102
Defocus ..... 104
Tilt ..... 104
Quadratic Piston Error ..... 104
Jones Matrix Form for the Wavefront Aberrations ..... 105
6. POLARIZATION ABERRATIONS ..... 106
Introduction ..... 106
Prior Work ..... 106
The Polarization Aberration Expansion ..... 109
Real, Imaginary and Complex Terms ..... 109
The Polarization Aberration Expansion ..... 110
Description of Terms ..... 110
Single Surface Aberrations ..... 112
Discussion of the Terms ..... 113
The Geometric Origin of Tilt and Piston ..... 114
Scalar Polarization Aberrations ..... 117
Scalar Constant Piston Terms ..... 117
Scalar Defocus Terms ..... 120
Scalar Tilt Terms ..... 124

## TABLE OF CONTENTS--Continued

Page
Scalar Quadratic Piston Terms ..... 125
Vector Polarization Aberrations ..... 126
Vector Constant Piston Terms ..... 126
Vector Defocus Terms ..... 129
Vector Tilt Terms ..... 130
Vector Quadratic Piston Terms ..... 133
7. CALCULATION OF ABERRATION COEFFICIENTS ..... 134
Introduction ..... 134
Single Surface Aberrations for Amplitude Transmission Relations ..... 134
Example - Spherical Metal Mirror ..... 136
Object at Center of Cūrvaiure ..... 139
Polarization Aberrations for Cascaded Surfaces ..... 140
Paraxial Approximation ..... 140
System Calculation ..... 140
8. ALTERNATE POLARIZATION ABERRATIGN EXPANSIONS ..... 144
Introduction ..... 144
The C Vector Aberration Expansion ..... 145
The Exponential Phase Polarization Aberration Expansion ..... 146
The E Vector ..... 149
Order Independence ..... 149
Relation to the $\mathbf{C}$ Vector ..... 149
Nondiagonalizability ..... 151
Multiplication of $E$ Vectors ..... 152
Divergence ..... 152
The Exponential Polarization Aberration Expansion ..... 153
9. FINAL COMMENTS ..... 154
Introduction ..... 154
Strong Polarizers ..... 155
Order Dependence ..... 156
Names of Polarization Aberrations ..... 157
Polarization Ray Tracing ..... 157
Comparison of Polarization Aberrations and
Polarization Ray Tracing ..... 158
Appendix A: Notation ..... 161
Appendix B: A Comparison of the Jones and Mueller Calculus ..... 168
The Mueller Calculus ..... 168
Depolarization ..... 169
What the Jones and Mueller Calculus Describe ..... 169

## TABLE OF COPSTENTS--Continued

Page
Measuring the Jones and Mueller Parameters ..... 170
Mathematical Complexity ..... 171
Characterizing Depolarization ..... 172
Relationship between the Jones Matrix and the Mueller Matrix ..... 173
Summary ..... 174
Appendix C: The Meaning of the Coefficients of the C Vector ..... 177
Absorption ..... 177
Propagation, Phase Delay ..... 178
Polarization ..... 178
Retardance ..... 180
Appendix D: Paraxial Skew Rays ..... 184
Introduction ..... 184
Coordinate System ..... 184
Subscripts ..... 185
Meridional Rays, y-z Plane ..... 185
Paraxial Skew Rays ..... 186
Angle and Plane of Incidence for Radially Symmetric
Systems ..... 187
Paraxial Relations for the Angle of Incidence ..... 189
Refraction ..... 192
Reflection ..... 193
Appendix E: Selected Bibliography ..... 194
References ..... 206

## LIST OF TABLES

Table Page
1 The Eight Forms of Polarization Behavior ..... 4
2 Polarizing Optical Elements Listed Roughly by "Strength" ..... 6
3 Commercially Available Optical Design Programs ..... 16
4 Basis Jones Vectors ..... 33
5 The Identity Matrix and the Pauli Spin Matrices ..... 38
6 Jones Matrices and C Vectors for Ideal Polarizers ..... 41
7 Eigenvalues and Eigenvectors of the Pauli Spin Matrices ..... 45
8 Polarization States Transmitted by the Basis Pclarizers ..... 47
9 Interpretation of the C Vector Elements ..... 52
10 Notation for Chapter 4 ..... 90
11 Wavefront Aberrations ..... 103
12 Transforming Jones Matrices to Mueller Matrices ..... 175

## LIST OF ILLUSTRATIONS

Figure Page
1 Linear Polarization of a Spherical Refracting Interface ..... 11
2 The Imaging Transformation ..... 13
3 Incident, Reflected and Refracted Waves at an Interface ..... 65
4 Power Transmission of Dielectric Interfaces ..... 72
5 A Single Layer Thin Film in Transmission ..... 73
6 Power Transmission of a Quarter Wave Thin Film ..... 76
7 The Phase Change of a Quarter Wave Thin Film ..... 78
8 Transmission of an Antireflection Coating ..... 79
9 Phase Change and Retardance of an Antireflection Coating ..... 80
10 Coordinate System for Paraxial Rays ..... 98
11 Paraxial Angles of Incidence ..... 100
12 On and Off-Axis Angle of Incidence ..... 115
13 The Off-Axis Angle of Incidence Squared ..... 116
14 The Scalar Aberrations. Defocus, Tilt and Piston ..... 118
15 Circular Polarization Defocus, Tilt and Piston ..... 119
16 Optical Path Length through a Lens ..... 122
17 Amplitude Transmittance of Weakly Absorbing Lens ..... 123
18 The Second Order Linear Polarization Aberrations ..... 127
19 The Second Order Diagonal Polarization Aberrations ..... 128
20 The Angle of Incidence Squared for an On-Axis Object ..... 131
21 Linear Polarization Defocus ..... 132
22 Spherical Mirror ..... 137
23 Y-Y Bar Parameters for an Interface ..... $19!$


#### Abstract

Polarization aberrations are the variations of amplitude, phase, polarization and relardance associated with ray paths through optica! systems. This dissertation develops methods for calculating the polarization aberrations of radially symmetric systems of weak polarizers, systems like lenses. telescopes and microscopes. The instrumental polarization in these systems arises from weak polarization effects occurring near normal incidence at glass. metal and thin film coated interfaces.

Polarized light and polarizers are treated using the Jones calculus. Weak polarizers, opticai elements with small polarization effecis, are ireated by expanding the Fresnel equations and thin film equations into a Taylor series. Methods are given for calculating the Taylor series coefficients for an multilayer coated interface whose polarization performance is known, for example from a thin film design program. Equations are derived for the propagation of polarized light through optical systems. Weak polarizers are shown to be very weakiy order dependent; this greatly facilitates the calculation of the effect of a sequence of weak polarizers. The dominant tersns are order independent polarization terms which are readily calculated. The order dependent portion can be systematically evaluated as higher order terms.

The instrumental polarization, being a function of angle of incidence, is different for different rays through the system. Thus an optical system is a


spatially varying polarizer. The instrumental polarization associated with a single surface is often well approximated as a "parabolic" polarizer. The instrumental polarization function is calculated as a Taylor series Jones matrix abcut the optical axis as a function of object and pupil coordinates. The resulting spatial variations of the instrumental polarization function bear a strong resemblance to the wavefront aberrations, since both arise from fundamental geometrical considerations. In particular, there are terms in the weak linear polarization and in the weak retardance of radially symmetric systems which strongly resemble defocus, tilt and piston error. A polarization aberration expansion is defined to secoũl ourder in the objeci and pupii coordinaies. A method is derived for calculating the polarization aberration coefficients for a sequence of radially symmetric surfaces from the Taylor series representation of the polarization associated with the individual interfaces.

## CHAPTER 1

## INTRODUCTION

This dissertation is a theoretical study of the propagation of polarized light through optical systems. Expressions for the instrumental polarization of radially symmetric optical systems of lenses, mirrors and coatings in the paraxial region have been derived. This instrumental polarization has been found to have a mathematical form similar to the wavefront aberrations of geometrical optics. A set of functions have been derived to characierize the instrumentiai poiarization of symmetric optical systems. These functions have been named the "polarization aberrations". The polarization aberrations are a generalization of the wavefront aberrations and include the wavefront aberrations as a subset. What is new and unique about the present work is the emphasis on determining the variation of amplitude and polarization in the transmitted beam as functions of object and pupil coordinates. This is a relatively unexplored area of optical design with little prior research.

Polarizers are often incorporated into optical systems to control the state of polarization of light through the system. But all optical interfaces, including reflecting and refracting interfaces, have intrinsic polarizing properties. Since the polarization properties of optical interfaces vary with the angle of incidence of the light, nonplanar optical surfaces become spatially varying polarizers which change the amplitude and polarization of transmitted light in a complex fashion. It is the characterization of optical systems as polarizers, typically weak polarizers, and the calculation of the magnitude of these polarization variations in optical systems that this dissertation treats.

This dissertation derives the relationships needed to calculate the polarization aberrations for the class of radially symmetric optical systems of lenses, mirrors and coatings.

## Polarized Light

Light is a transverse electromagnetic wave. The electric and magnetic fields associated with the optical disturbance are perpendicular to the local direction of propagation, the Poynting vector. Thus, light is a vector wave; the associated fields have two degrees of freedom associated with the two directions orthogenal to the Poynting vector. Polarizatioun refers io ine properies of iignt associated with these two degrees of freedom. Electromagnetic waves share this property, polarization, with other vector fields such as elastic and spin waves in solids.

The mathematics necessary to describe the polarization of light is contained in Chapters Two and Three.

## Polarization Elements

Polarization elements are optical elements which change the polarization state of light. Polarization elements can be grouped into several broad and overlapping categories: polarizers, retarders, linear elements, elliptical elements, circular elements and depolarizers. Polarizers, such as dichroic "polaroid" sheets and polarizing prisms, preferentially transmit certain polarization states. Retarders, such as birefringent plates, introduce a phase delay between different polarization states. Linear elements have linear "eigenpolarizations" while elliptical and circular elements have eiliptical or circular eigenpolarizations. Depolarizers randomly alter the phase information of light causing an irreversible mixing of polarization states.

At this point a word shortage is encountered. The word polarizer is used for elements which produce amplitude differences when acting upon different polarization states. Likewise, retarders produce a phase difference acting upon different polarization states. However elements which produce both amplitude and phase differences, such as most coatings, are also generally referred to as polarizers. The words polarization element are used here to make the distinction. Clearly another word for the amplitude polarization elements would be helpful but the literature does not make any clear distinctions. The word dichroism is close but it refers specifically to the material property of a polarization dependent absorption coefficient. Dichroism is not strictly appropriate for coatings which produce amplitude differences by interference. This shortage of words makes the words polarization and polarizer somewhat ambiguous and only serves to needlessly complicate this subject.

## The Eight Forms of Polarization Behavior

There are only eight forms of polarization behavior associated with a nonscattering polarizer. These are listed in Table 1. These eight forms correspond to the eight degrees of freedom in the Jones matrix's four complex elements. The instrumental polarization function contains the linear polarization, retardance, amplitude and other polarizing terms associated with ray paths as well as the wavefront aberration (which is all conventional ray tracing calculates).

## Instrumental and Residual Polarization

Instrumental polarization is the polarization and retardance intrinsically associated with the optical elements in an optical system. Instrumental polarization

## TABLE 1

## THE EIGHT FORMS OF POLARIZATION BEHAVIOR

| 1. | Amplitude. |
| :--- | :--- |
| 2. | Phase (Wavefront Aberration), |
| 3.4. |  |
| 5.6. | LinearPolarization (Magnitude and Orientation), |
| 7. |  |
| 8. |  |

arises from lenses, mirrors, coatings, diffractive optical elements and crystals as well as from polarizers and retarders. The instrumental polarization can be divided into contributions from optical elements used specifically for polarization control, polarizers and retarders, and the "residual polarization." which is associated with lenses, mirrors, coatings, gratings, holograms and other optical elements. Residual polarization will be defined as polarization contributions from optical components not specifically intended as polarizing elements. Residual polarization is undesirable since any polarization or retardance present can couple light between orthogonal polarization states. Residual polarization might be compared to wavefront aberration because both interfere with the measurement of optical fields and rejuce the image forming potential of the optical system.

## Sources of Instrunefintal Polarization

All optical interfaces display some polarization effects at nonnormal incidence. The $\vec{E}_{s}$ electric field component drives electrons tangential to the surface while the $\vec{E}_{\mathbf{p}}$ component tries to force electrons into the metal and then draw them out. The response of the interface to these two stimuli is different. For example, when reflecting from a metal, the $s$ component is reflected more efficiently than the p component; the interface is a weak linear polarizer for non-normal incidence light. In addition, there is a phase difference upon reflection between the $s$ and $p$ components so the interface is also a weak retarder.

Table 2 is a list of polarizing optical elements with the more strongly polarizing elements towards the top of the list and weaker polarizers towards the bottom.

## TABLE 2

## POLARIZING OPTICAL ELEMENTS LISTED ROUGHLY BY "STRENGTH"

| Polarizers | Holograms |
| :--- | :--- |
| Retarders | Fold Mirrors |
| Electro-Optic Crystals | Dichroic Filters |
| Optical Fibers | Bandpass Filters |
| Dichroic Crystals | Mirrors |
| Beamsplitters | Lenses |
| Waveguides | Antireflection Coatings |
| Grazing Incidence Mirrors | Gradient Index Media |
| Strain Birefringent Media |  |

## Systems Sensitive to Instrumental Polarization

Instrumental polarization is the polarization and retardance intrinsically associated with the optical elements in an optical system. Instrumental polarization arises from lenses, mirrors, coatings, diffractive optical elements and crystals as well as from polarizers and retarders. The instrumental polarization can be divided into contritutions from optical elements used specifically for polarization control, polarizers and retarders, and the "residual polarization," which is associated with lenses, mirrors, coatings, gratings, holograms and other optical elements. Residual polarization will be defined as polarization contributions from optical components not specifically intended as polarizing elements.

Many optical systems are intended to transmit all polarization states equally. The collection optics for a radiometer used in remote sensing should have equal transmittance for any incident polarization state. Since natural scenes are usually partially polarized, the radiometer will not be biased toward a particular polarization. Similarly, the optics in a camera or copy machine do not benefit from having a different transmittance for different polarizations; this can only represent an additional loss of light with no advantages for the system. A large class of systems work better if they are free of linear polarization, retardance and depolarization.

In instruments used for the measurement of the polarization state of light. such as polarimeters and ellipsometers, any residual polarization leads to inaccurate polarization measurements. To mitigate this degrading factor, the lenses and coatings used in polarimeters should be chosen very carefully to minimize the instrumental polarization.

In interferometers, any light coupled into the orthogonal state does not produce interference fringes with the primary polarization component. The residual polarization leads to loss of fringe visibility. The polarization effects (linear or circular polarization) produce variations of exit pupil brightness which depend on the incident polarization state. These variations make it difficult to match the intensities of the sample and reference beams across the pupil. Further, the match is different for different input polarizations. The retardance effects (phase polarization) cause polarization dependent wavefronts. Different input polarization states produce different interferograms.

Residual polarization effects are most troublesome in advanced interferometric instruments such as laser radars, optical phased arrays and optical signal processors, systems where as much information as possible is extracted from interference patterns. In these systems, signal to noise is reduced in proportion to the amount of residual polarization; it is a first order effect. Similar first order problems occur in optical communication systems utilizing local oscillators for heterodyne detection.

Instrumental polarization is a substantial effect in systems with large angles of incidence, such as those coniaining holograms, fold mirrors, diffraction gratings or grazing incidence systems.

Systems operating over a broad spectral band often encounter large polarization effects from coatings. particularly at the edges of the spectral bandpass. Systems with widely separated laser wavelengths typically show enhanced instrumental polarization at one or both wavelengths. Strong polarization effects frequently occur at the transmission edges of band-pass filters, edge filters and dichroic filters.

High precision radiometers, such as those used for absolute measurements to tenths of a percent, are instrumental polarization sensitive and require very high quality coatings. One percent of instrumental polarization results in the radiometer recording 1 watt of incident optical power as anywhere between 0.99 watt and 1.01 watt depending upon the incident polarization state. In remote sensing systems, there is no control over the polarization state of the incident light.

## Polarization Aberrations

The causes of instrumental polarization in optical systems are the polarization produced by non-normal incidence at the optical interfaces and the polarization produced propagating through polarizing media. Since each ray takes a different path through the system with its own angles of incidence and planes of incidence, each ray in general experiences a different change in its state of polarization. This residual polarization varies with wavelength, object coordinates and pupil coordinates. "Polarization aberrations" will be defined as variations of the amplitude, phase and polarization of an optical wavefront across the exit pupil of an optical system and the dependence of these variations on wavelength and object coordinate. The polarization aberrations are extensions of the wavefront aberrations. Since the polarization aberrations encompass amplitude and polarization variations they provide a more complete characterization of the electromagnetic fields transmitted by an optical system.

The method of polarization aberrations combines elements of several different optical calculation methods into one procedure. Its objective is the calculation of the instrumental polarization function of an optical system. The method of polarization aberrations supplements the aberration equations of geometrical optics (Born and

Wolf 1975, Smith, Kingslake 1978) with thin film calculations (Born and Wolf 1975, Macleod ) and crystal optics calculations (Born and Wolf 1975, Yariv and Yeh 1974.) Using polarization aberrations a more complete picture of the wavefront transmitted by an optical system is obtained. In particular, since thin films and anisotropic crystals are polarizing, the transmitted amplitude and wavefront are usually a function of the incident polarization state. Specifically, the amplitude as well as the amounts of defocus, spherical aberration, astigmatism and other aberrations of a system vary with the incident polarization state. These are the effects that the polarization aberration method calculates.

One example of this analysis is that optical interfaces behave as a spatially varying weak linear polarizers and spatially varying weak linear retarders. For example, an uncoated spherical lens surface interacting with an on axis spherical wave has an associated linear polarization due to the Fresnel equations whose polarization axes are oriented radially and whose linear polarization magnitude increases quadratically with pupil radius. Figure la depicts this linear polarization aberration (spatial variation) across the lens aperture. Each line represents the weak linear polarization the beam experiences at that part of the surface. The lines are drawn parallel to the axis of the local weak linear polarization and the lengths of the lines are proportional to the polarizance, the degree of polarization produced when the incident light is unpolarized. This residual linear polarization is vertically oriented along the y axis and horizontally oriented along the x axis. This weak spatially varying polarizer is an example of a polarization aberration.

Continuing this example, Figure 1 b shows the effect of the lens surface on a uniform intensity beam linear polarized along the $y$ axis. The lines show the


## Figure ! Residual Polarization of a Spherical Refracang Interface

a. The magnitude and orientation of the instrumental linear polarization across a spherical interface is zapresented by the length and orientation of lines.
b. The transmitted polarization state for vertical liseat polarized incident light is represented by the length and angle of the arrows.
intensity and orientation of the polarization of the beam after being refracted by the uncoated interface. Along the $y$ axis, where the beam is aligned with the instrumental linear polarization, the beam becoms brighter towards the top and bottom of the aperture. This increase in transmission occurs because the light is approaching Brewster's angle. Along the x axis the light is orthogonal to the residual linear polarization and becomes less intense towards the edge. At $\pm 45$ degrees, the light is polarized at an angle to the residual linear polarization and has its polarization axis rotated; thus light is coupled into the orthogonal polarization state upon refraction.

## Optical Design

Electromagnetic waves are characterized by their amplitude, phase, polarization and cohereace. The amplitude describes the intensity distribution in space. The phase describes the shape of the wavefronts. The polarization describes the direction of oscillation of the electromagnetic waves. The coherence collectively describes the statistical properties of the amplitude, phase and polarization. For example, the degree of polarization is a measure of the "angular coherence".

Optical design, as an activity, is the process of selecting a set of optical surfaces on which to perform a desired transformation of the optical fields incident upon the system. The most important transformation is imaging. depicted in Figure 2. Spherical waves emanating from a region called the object are to be transformed into converging spherical waves. At the centers of curvature of these spherical waves, the image, a likeness of the object, is formed, often on a screen or a piece of film. Other importaat transformations include coupling light into an aperture, providing a uniform or specified illamination, interfering two or more beams of

## .



Figure 2 The Imaging Transformation
An ideal imaging system transforms diverging spherical waves from the object into spherical waves converging towards image.
light, maximizing reflectance or transmittance, modulating or scanning a beam, and providing a spectral decomposition of the incident light. The materials available include: lenses, mirrors, prisms, diffraction gratings, optical fibers, polarizers. diffusers, thin film coatings, waveguides and crystals.

Optical design then consists of calculating or mathematically simulating the behavior of electromagnetic waves through an assembly of optical elements and choosing the elements in suicin a way as to optimize optical performance.

Optical theories can be arranged in a hierarchy of increasing complexity and sophistication: geometrical optics, wave optics and diffraction theory, quantum optics and quantum electrodynamics. Most optical engineering problems, such as imaging, can be adequately simulated using geometrical optics with some wave optics calculations to evaluate the diffraction performance. The fundamental operation is ray tracing, calculating the path of the normal to a wavefront through the system by the repeated application of Snell's law, the law of reflection and the diffraction grating law. By tracing a collection of rays from a single object point, the location of an image, a relatively small region in space where the rays converge and nearly intersect, can be determined. One important measure of the image quality calculated by ray tracing is the variation of the optical path difference, usually calculated at the exit pupil of the system, for different rays from a given object point. Ray tracing will also determine the locations of ray intercepts at the image plane, the ray aberrations, and the optical path difference, of relative phase, along these ray paths. Ray tracing is the repeated application of several relatively simple physical laws to relatively complex arrangements of optical surfaces. Thus optical design has evolved into the use of elaborate computer programs to perform the enormous amounts of ray
tracing and associated data reduction necessary for the design of optical systems. Table 3 contains a list of commercially available ray tracing programs commonly used for optical design.

Of the parameters necessary to describe the optical fields; amplitude, phase, polarization and coherence; ray tracing only calculates the phase. By calculating ray paths and optical path lengths through the system, ray tracing determines the precise shape of the wavefront, but does not determine the amplitude or polarization. Centuries of experience have shown that the phase (as determined from the optical path length) is the most critical of these parameters in the design process. Born and Wolf (1975, chap.9) give an extended discussion of the effect of phase errors on optical images. The variation of optical path length along the ray paths forming an image must be held to a small fraction of a wavelength or the structure of the diffraction pattern will suffer major degradation. Thus the majority of effort in optical design has been spent controlling the shape of the transmitted wavefront. The optical designer's main task is to determine an optical system which satisfies specifications on the wavefront quality for a range of wavelengths and field positions. All of the computer programs listed in Table 3 perform calculations to accurately determine the optical path differences associated with images and will calculate the effect of these phase differences on the diffraction image. None of these programs will calculate the instrumental polarization and its variation for different rays or the effect of these polarization aberrations on the image. This is understandable since these effects are, for the most part. small and have not yet been systematically explored.

TABLE 3

SOME COMMERCIALLY AVAILABLE OPTICAL DESIGN RAY TRACING PROGRAMS

| NAME | VENDOR | ADDRESS |
| :--- | :--- | :--- |
| ACCOS V | Scientific Calculations | Fishers, NY |
| CODE V | Optical Research Assoc. | Pasadena, CA |
| Cool/Genie | Genesee Computer Center | Rochester, NY |
| Kidger Program | Kidger Optics Limited | Crowborough, UK |
| OSLO | Sinclair Optics | Fairport, NY |
| Synopsis | Optical Systems Design | Medford, MA |

## Thin Films

Vacuum deposited thin films are used on most optical surfaces to control the amount of transmission and reflection. These thin films are usually less than the wavelength of light in thickness. Being so very thin, the effect of the films on ray paths are accurately modeled by treating the films as having parallel surfaces which contour the substrates on which they are deposited. Due to the closely spaced parallel surfaces, thin films have negligible influence on the ray paths through the system and are generally ignored when simulating a system by ray tracing. These coatings principally affect the amplitude and polarization of the ray and have much less effect on the optical path difference. This division, with the optical surfaces governing the ray paths and the thin film coatings governing the amplitude and transmission. allows the optical system design problem to neatly deccuple into two separate problems, lens design and coating design. The wavefiont performance and image quality of the system is calculated by a lens designer using a ray tracing optical design program. The amplitude and polarization calculations at individual surfaces are performed by a coating designer using a thin film design program. This division of labor usually produces a loosely coupled communication channel between the optical and coating designers. Typically, the coating designers are only given the wavelength range and angle of incidence range for a surface and asked to design coatings to maximize transmittance or reflectance under these broad conditions. One reason this loosely coupled design process has worked so well is because of a fortuitous circumstance. The coatings designed to optimize the transmittance or reflectance at an interface have usually greatly reduced the amplitude and polarization variations and thus reduced the polarization aberrations at the interface
as well. For example, a quarter wave magnesium fluoride antireflection coating on glass typically reduces reflection losses at the design wavelength by a factor of four, and reduces the instrumental polarization by a still greater factor (see Chapter 4). This fortuitous circumstance has allowed lens and coating design to remain uncoupled. Thus instrumental polarization is usually ignored as a higher order effect. But it is not sufficient to design thin film coatings in isolation from the lens design when amplitude and polarization performance is demanding. Very low instrumental polarization is required for a new and more complex generation of optical systems such as laser radars, precision remote sensing spectropolarimeters, grazing incidence optics and phased array optical systems. Calculating the instrumental polarization requires performing thin film and crystal optics calculations during the ray tracing process. This idea is not new, but its implementation is complex enough to have delayed the integration of these two uranches of optical design until specifications required it.

Control of the wavefront aberrations has reached a state of refinement such that some systems are limited by instrumental polarization. Using computer controlled polishing and phase measuring interferometry, the best optical surfaces are now fabricated to sphericity of greater than 0.005 waves root mean square (rms). Several commercial systems, such is the Perkin Elmer Micralign system, are routinely produced with a system wavefront quality of better than 0.025 waves rms. Relatively small amounts of polarization aberration are needed to produce a change in tise Strehl ratio equal to the degradation caused by this very small residual wavefront aberration. For such carefully designed and fabricated systems, further improvements in image quality will come more easily from an understanding and
control of the instrumental polarization than from further reduction of the wavefront aberrations. Since interferometric and multiple beam systems are far more susceptible than imaging systems to amplitude and polarization errors. there is a need to perform the design of these systems and their coatings in a more tightly coupled manner by using procedures such as polarization ray tracing and polarization aberrations to simulate the final instrumental polarization performance.

## The Wavefront Aberration Function

There are two principal methods for determining the wavafront transmitted by an optical system, aberration theory and ray tracing. Both calculate the optical path lengths for ray paths through an optical system. One principal result of these methods is the determination of the wavefront aberration function (or optical path difference function) in the exit pupil for a specific object coordinate and wavelength. The wavefront aberration function is expressed as $W(\vec{h}, \vec{\rho})$ where $\vec{h}$ and $\vec{\rho}$ are the object and pupil coordinates. The wavefront aberration function is a dimensionless quantity (expressed in wavelengths) which describes the shape of the transmitted wavefront. It does not contain information about the transmission of the optical system or the polarization state of the transmitted light.

When optical design programs utilize ray tracing information as input to diffraction calculations, most make the "default assumption of geometrical optics," that the transmitted wavefront has uniform amplitude and constant polarization state across the exit pupil. The default assumption of geometrical optics is only valid as long the amplitude variations and wavefront differences between the two polarization components are small enough to be negligible for the problem at hand. In the absence of further information about the transmitter wavefront the default
assumption is the safe and practical assumption. For the majority of optical systems it is an excellent assumption. Polarization aberrations is a method for designing and analyzing systems where this assumption is inadequate.

The default assumption of geometrical optics combined with the results of ray tracing leads to an expression for the time independent electromagnetic field complex amplitude $\alpha(\overrightarrow{\mathrm{h}}, \vec{\rho})$ on a reference sphere in the exit pupil of

$$
\mathrm{a}(\overrightarrow{\mathrm{~h}}, \vec{\rho})=\hat{\mathrm{p}} \mathrm{a}_{0} \mathrm{e}^{\mathrm{j} 2 \pi \mathrm{~W}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) / \lambda} .
$$

Here, $\hat{p}$ is the polarization state, $a_{0}$ is the real amplituce of the field, $\lambda$ is the wavelength and $\mathrm{W}(\overrightarrow{\mathrm{h}}, \vec{\rho})$ is the wavefront in the exit pupil (coordinates $\vec{\rho}$ ) for object coordinate $\overrightarrow{\mathrm{h}}$. Due to the default assumption of geometrical optics, the polarization state $\hat{\mathbf{p}}$ and the real amplitude $\mathrm{a}_{0}$ are constants, independent of $\overrightarrow{\mathrm{h}}$ and $\vec{\rho}$.

## The Instrumental Polarization Function

A more general expression for the time independent electromagnetic field complex amplitude which encompasses amplitude variations and polarization variations is

$$
a(\overrightarrow{\mathrm{~h}}, \vec{\rho})=\hat{\mathrm{p}} \mathrm{a}_{\mathrm{p}}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) \mathrm{e}^{\mathrm{j} 2 \pi \mathrm{~W}_{\mathrm{p}}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) / \lambda}+\hat{\mathrm{q}} \mathrm{a}_{\mathrm{q}}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) \mathrm{e}^{\mathrm{j} 2 \pi \mathrm{~W}_{\mathrm{q}}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) / \lambda}
$$

Here $\hat{p}$ and $\hat{q}$ are any two orthogonal polarization states such as horizontal and vertical linearly polarized light or left and right circularly polarized light. The real amplitudes of the two polarization components, $a_{p}(\vec{h}, \vec{\rho})$ and $a_{q}(\vec{h}, \vec{\rho})$, are now functions of the pupil and object coorainates. Likewise $W_{p}(\vec{h}, \vec{\rho})$ and $W_{q}(\vec{h}, \vec{\rho})$ are the two wavefronts associated with the two polarization components.

The functions $\mathrm{a}_{\mathrm{p}}(\overrightarrow{\mathrm{h}}, \vec{\rho}), \mathrm{a}_{\mathrm{q}}(\overrightarrow{\mathrm{h}}, \vec{\rho}), \mathrm{W}_{\mathrm{p}}(\overrightarrow{\mathrm{h}}, \vec{\rho})$ and $\mathrm{W}_{\mathrm{q}}(\overrightarrow{\mathrm{h}}, \vec{\rho})$ cannot be determined by conventional ray tracing and aberration theory since they are functions of the
polarization and retardance effects associated with the optical elements.
The importance of considering the two polarization components separately arises since orthogonal components do not produce a stationary interference pattern. The two polarization components form essentially two separate diffraction patterns which are added incoherently to form the point spread function. If $a(\vec{h}, \vec{\rho})$ and $W(\vec{i}, \vec{\rho})$ are equal for the two polarization components, then these separate polarization diffraction patterns are equal and the polarization effects are trivial. In the presence of weakly polarizing or weakly retarding components in the optical system such as optical coatings, a small fraction of the light in one polarization state is transferred into the orthogonal state which now has a different amplitude and phase. This coupled light forms a second "ghost" diffraction pattern lurking alongside the primary diffraction pattern formed by the principal polarization sate (Kubota and Inoue 1959, Fainman and Shamir 1984.)

This transfer of a fraction of the light from one polarization state into another has frequently and incorrectly been labeled "depolarization." Properly stated, depolarization is the coupling of polarized light into unpolarized light. Depolarization is related to scattering. A loss of degree of polarization is associated with all depolarizing processes. Since no loss of degree of polarization is implied by the present process, only a change of polarization state, the correct term is "polarization coupling."

## The Jones Calculus Representation of the Instrumental Polarization Function

The instrumental polarization function for an optical system is derived using the Jones calculus. The following analysis includes the effects of reflections. refractions. thin films and anisotropic media and neglects all scattering and
depolarization mechanisms. The time independent electromagnetic field complex amplitude expressed in Jones calculus notation is

$$
\vec{J}(\overrightarrow{\mathrm{~h}}, \vec{\rho})=\left[\begin{array}{c}
a_{x}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) \mathrm{e}^{\mathrm{j} 2 \pi \mathrm{~W}_{x}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) / \lambda} \\
\mathrm{a}_{y}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) \mathrm{e}^{\mathrm{j} 2 \pi \mathrm{~W}_{\mathrm{y}}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) / \lambda}
\end{array}\right] .
$$

The Jones vector function for the complex amplitude $\vec{J}$ describes the transmitted complex amplitude for a specified input polarization state. A more complete description of the instrumental polarization describes the transmitted amplitude for arbitrary input polarization states. The general expression for the "instrumental polarization function" along ray paths through the system as a Jones matrix $\mathbf{J}$ of the object and pupil coordinates is

$$
\mathbf{J}(\overrightarrow{\mathrm{h}}, \vec{\rho})=\left[\begin{array}{ll}
\mathrm{j}_{11}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) & \mathrm{j}_{12}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) \\
\mathrm{j}_{21}(\overrightarrow{\mathrm{~h}}, \vec{\rho}) & \mathrm{j}_{22}(\overrightarrow{\mathrm{~h}}, \vec{\rho})
\end{array}\right]
$$

Given the instrumental polarization function in Jones matrix form, the transmitted amplitude in the exit pupil is then known for arbitrary ray paths and arbitrary input polarization states.

The primary objective of this work the calculation of the instrumental polarization function $\mathbf{J}(\overrightarrow{\mathrm{h}}, \vec{\rho})$.

## Jones Matrix Form fer Interface Polarization

Reflection and refraction at a homogeneous and isotropic interface, such as a metallic reflecting surface, a dielectric lens surface, or a thin film optical coating on a surface, is characterized by amplitude transmission coefficients for the $s$ and $p$
components, $a_{s}$ and $a_{p}$. The amplitude transmission coefficients relate the electric field components before and after the interface:

$$
\begin{aligned}
& \vec{E}_{s}^{\prime}=a_{s} \vec{E}_{s}=\rho_{s} e^{j \delta_{s}} \vec{E}_{s} \\
& \vec{E}_{p}^{\prime}=a_{p} \vec{E}_{p}=\rho_{p} e^{j \delta_{p}} \vec{E}_{p}
\end{aligned}
$$

A measure of the linear polarization associated with an interface at a particular wavelength and angle of incidence is the polarizance (Shurcliff 1961.)

$$
\frac{\rho_{S}^{2}-\rho_{p}^{2}}{\rho_{S}^{2}+\rho_{p}^{2}}
$$

Likewise

$$
\delta=\delta_{s}-\delta_{p}
$$

is the linear retardance for the interface in radians. The coefficients, $\mathrm{a}_{\mathrm{s}^{\circ}} \mathrm{a}_{\mathrm{s}^{\prime}} \rho_{\mathrm{s}^{\prime}} \rho_{\mathrm{s}^{\prime}}$ $\delta_{\mathbf{s}^{\prime}}$ and $\delta_{\mathbf{S}^{\prime}}$ can be obtained using thin film design programs for a large variety of interfaces.

From the amplitude transmission coefficients for a nonscattering homogeneous and isotropic interface a Jones matrix can be written for a ray at an interface. The local ray information required is the angle of incidence of the ray, $i$, and the orientation of the plane of incidence, $\theta$. In s-p coordinates the Jones matrix is

$$
J_{s} p(i)=\left[\begin{array}{cc}
a_{s}(i) & 0 \\
0 & a_{p}(i)
\end{array}\right]
$$

This is the Jones matrix for a linear polarizer aligned with a linear retarder. This matrix must be rotated into the system $x-y-z$ coordinates before multiplying it with the Jones matrices for the other optical elements. Using the rotation operator $\mathbf{R}(\theta)$ for the Jones calculus,

$$
\mathbf{J}(\mathrm{i}, \theta)=\mathbf{R}(\theta) \mathbf{J}(\mathrm{i}) \mathbf{R}(-\theta)=\left[\begin{array}{rr}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{array}\right] \mathbf{J}(\mathrm{i})\left[\begin{array}{r}
\cos \theta \\
\sin \theta \\
-\sin \theta \\
\cos \theta
\end{array}\right] .
$$

More complex forms for the Jones matrix describe interfaces which are not isotropic such as diffraction gratings, holograms, and anisotropic crystals.

## The Jones Matrix for Propagation

The Jones matrix for propagation through a length $:$ of nonpolarizing material of refractive index $n$ is

$$
J(l, n)=e^{j 2 \pi n \ell / \lambda}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]
$$

Absorbing materials are treated though the use of a complex n. Related Jones matrix equations for propagation through anisotropic media and optical modulators are collected in a recent textioook by Yariv and Yeh (1984.)

## The Jones Matrix for Ray Paths Through Optical Systems

Next the calculation of the Jones matrix associated with an arbitrary ray paith through an optical system is presented. The ray originates at object coordinate $\overrightarrow{\mathrm{h}}_{0}$ and enters the system at entrance pupil coordinate $\vec{\rho}_{0}$. Let $Q$ be the number of optical interfaces in the system. The Jones matrix associated with the ray at surface $\mathbf{q}$ is $\mathbf{J}_{\mathbf{q}}$ while the Jones matrix for propagation from interface $\mathbf{q}$ to $\mathbf{q + 1}$ is $\mathbf{J}_{\mathbf{q}+1, \mathbf{q}}$ The Jones matrix for the ray from object space ( $q=1$ ) to image space $(q=Q$ ) is

$$
J\left[\vec{h}_{0,} \vec{\rho}_{0}\right]=J_{Q} \mathbf{J}_{Q, Q-1} J_{Q-1} J_{Q-1, Q-2} \cdots J_{2} J_{2,1} J_{1}=\prod_{q=Q,-1}^{1} J_{q} J_{q, q-1}
$$

This Jones matrix is a single point in the instrumental polarization function.

There are two methods of progressing from this equation to the full instrumental polarization function, polarization aberrations and polarization ray tracing. Polarization aberrations provide approximate forms for $J(\vec{h}, \vec{\rho}, \lambda)$ in functional form by expanding $J(\vec{h}, \vec{\rho})$ about $J\left[\vec{h}_{0}, \vec{\rho}_{0}\right]$ given a Taylor series description of the optical elements. Polarization ray tracing calculates exact values of $J(\vec{h}, \vec{\rho}, \lambda)$ numerically at a discrete set of $\vec{h}, \vec{\rho}$, and $\lambda$ using this equation. Chapters $5,6,7,8$ contain the polarization aberration method. Polarization ray tracing is briefly presented in Ciiapter 9. The present work concentrates on the polarization aberrations since they provide more insight into the underlying physics.

## Interpretation of the Instrumental Polarization Function

The instrumental polarization function provides an abundance of data about the optical performance of an system. Where the wavefront aberration function is a single valued function of pupil coordinates, the instrumental polarization function is an eight valued function, a spatially varying Jones matrix. How can all this information be assimilated? What does it mean?

The straightforward interpretation of the instrumental polarization function proceeds by decomposing it into the eight basic forms of polarization behavior. One part is the wavefront aberration function which is interpreted in the conventional fashion. Another part is the transmission of the system for unpolarized light. The linear polarization terms contain the difference in transmission for different polarization states. The retardance terms contain the differences in the transmitted wavefront present between different polarization states. Details on this decomposition are contained in Chapters 6 and 7 and Appendix C.

The phase of the wavefront contains the most important information about the image forming potential of the wavefront. Due to the retardance portion of the instrumental polarization function however, the system has a different wavefront for different input polarization states. These variations are easiest to interpret if label as the "phase," the value of the wavefront midway between the maximum and minimum phase values possible. Then the polarization dependent deviations from this "phase" are the retardances of the system. The maximum and minimum deviations in phase are assumed by the eigenpolarizations of the poiarizer and are readily determined from the Jones matrix.

Plots can be made of the variations of the polarization performance parameters in a format like "rimray plots", plots of the parameter versus pupil coordinate along the x and y axes for several field positions. These figures help reveal the form of the polarization aberration of the system. Another useful form of polarization data display is maps of parameter variation across the exit pupil in contour plot or hidden line format. Such plots of the polarization alerrations are contained in Chapter 7.

## CHAPTER 2

## THE JONES MATRIX AND C VECTOR FOR THE CHARACTERIZATION OF POLARIZATION

In this chapter, the mathematical description of polarized light and polarizers is developed. There are two principal methods of handling polarization problems, the Jones calculus and the Mueller calculus, both developed in Cambridge. Massachusetts in the 1940's, Jones (1941a), Parke (1949). The Jones calculus is more appropriate for the development of the polarization aberrations since it is an amplitude calculus while the Mueller calculus is an intensity calculus. Appendix B contains a comparison of the two calculi and a discussiou of depolarization.

In this chapter, a variant of the Jones calculus, the $\mathbf{C}$ vector, is introduced and developed. The $\mathbf{C}$ vector, which decomposes the Jones matrix into a sum of Pauli spin matrices, is especially useful for problems involving weak polarizers, such as the instrumental polarization of highly transmitting systems.

## The Jones Calculus

The Jones calculus is a mathematical formalism introduced by R. Clark Jones (1941a) of the Polaroid Corporation and Harvard University to treat problems involving polarized light and polarizers. The Jones calculus was fully developed in a series of papers by $\overline{\mathbf{R}}$. C. Jones entitled "A New Calculus for the Treatment of Optical Systems": Jones (1941a), Jones (1941b), Jones (1941c), Jones (1942), Jones (1947a), Jones (1947b), Jones (1948) and Jones (1956). This series remains one of the
best treatments of the Jones calculus and contains information on the Jones calculus not found elsewhere. Jones' series of papers have been reprinted in Polarized Light. Swindell (1975). Other introductions to the Jones calculus are found in: Azzam and Bashara (1977, Section 1.6), Clarke and Grainger (1971, Section 1.3), Gerrard and Burch (1975, Section IV.5), Hecht and Zajac (1974), Shurcliff (1962, Chap.8) and Theocaris and Gdoutos (1979, Section 4.3.2.)

The Jones calculus description of polarized light and polarizers uses the Jones vector for the description of polarized light and the Jones matrix to characterize the polarizing properties of an optical element.

## Polarization Elements

Polarization elements are optical elements which divide an optical beam into two parts and transmit those parts with a different transmission coefficient and a different phase. The two parts of the beam are referred to the eigenvectors or by the more descriptive term, "eigenpolarizations," Azzam and Bashara (1977. pg.97). The two eigenpolarizations are orthogonally polarized and are transmitted by the polarizer with no alteration of their polarization states; only the intensity and phase changes.

The term polarization elements is used to refer to both polarizers, such as the dichroic or prism types, which have a different transmittance for the two eigenpolarizations, and retarders which have equal transmittance but a different phase change for the eigenpolarizations. Shurcliff's book, Polarized Light, (1962), is the standard reference on the types of polarization elements, their definitions, parameters and properties.

## Notation

Appendix A contains a list of most of the notation used in this dissertation. All vectors are be denoted with arrows $(\rightarrow)$ except for normalized vectors which are denoted with carets ( $\mathcal{\wedge}$ ). All matrices are printed in boldface.

## The Jones Vector

The expression for a quasimonochromatic plane wave propagating parallel to the z axis is.

$$
\vec{E}(t)=\vec{E}_{x}(t)+\vec{E}_{y}(t) \text {. }
$$

where,

$$
\begin{gathered}
\vec{E}_{x}(t)=\hat{x} E_{0, x}(t) \cos \left[[\hat{k} z-\bar{\omega} t]+\epsilon_{x}(t)\right], \\
\text { and, } \\
\vec{E}_{y}(t)=\hat{y} E_{0, y}(t) \cos \left[[\hat{k} z-\bar{\omega} t]+\epsilon_{y}(t)\right] .
\end{gathered}
$$

Ir these expressions, $\vec{E}_{\mathbf{x}}(\mathrm{t})$ and $\vec{E}_{\mathbf{x}}(\mathrm{t})$ are the instantaneous scalar components of $\vec{E}(t)$; $\hat{x}$ and $\hat{y}$ are the unit vectors aiong the enordinate axes. $\hat{\mathrm{k}}$ is the unit wavevector. $\bar{\omega}$ is the mean frequency. $\epsilon_{x}(t)$ and $\epsilon_{y}(t)$ are the adjustments to the $x$ and $y$ phase as a function of time. For coherent light the variations of $\epsilon_{x}(t)$ and $\epsilon_{y}(t)$ are much less than one radian during an optical period and the beam can interfere with itself with a visibility of near one in a Michelson interferometer for optical path differences of many wavelengths.

The time dependent Jones vector is defined in terms of the electric field amplitudes as

$$
J(t)=\left[\begin{array}{l}
\vec{E}_{x}(t) \\
\vec{E}_{y}(t)
\end{array}\right] .
$$

The components of $\vec{J}(t)$ are the instantaneous components of $\vec{E}(t)$.
The normalized Jones vector $J$ is a time independent normalized vector where the components of $\vec{J}(t)$ have been divided by the incident electric field amplitude,

$$
J=\frac{J(t)}{\vec{E}_{0}(t)} .
$$

The normalized Jones vector is referred to as "the Jones vector" unless otherwise stated. Knowledge of $\vec{J}$ and $\vec{E}_{0}$ provides all the information necessary to reconstruct $\vec{E}(t)$ to within a constant phase factor. $\vec{J}$ is written as either a column or row vector depending on the context.

## The Coherent Addition of Optical Fields

The coherent addition of two light fields propagating along the z axis is

$$
\vec{E}(t)=\vec{E}_{1}(t)+\vec{E}_{2}(t) .
$$

Written in Jones vector notation, this becomes

$$
\vec{J}=\vec{J}_{1}+\vec{J}_{2} .
$$

The extension to the coherent addition of N light fields is trivial,

$$
J=\sum_{n=1}^{N} J_{n} .
$$

This additive property makes the Jones vector quite suitable for the formulation of polarization problems in interferometry and diffraction theory.

## Relation to Elliptical Polarization Parameters

When $E_{X}(t)$ and $E_{y}(t)$ are plotted on an $x-y$ graph, a figure is traced out which describes the polarization state of the light. If the light is polarized, the same shape repeats indefinitely. The most general shape then possible for this figure is an ellipse, and the parameters of this ellipse are referred to as the elliptical polarization parameters. For quasimonochromatic completely polarized light, the shape of the ellipse remains fixed but the speed with which it is traced varies slightly as the wavelength drifts. For almost completely polarized light, the shape of the ellipse changes as the polarization state drifts. For unpolarized light, locally, the curve is an ellipse, but it has rapidly and randomly changing parameters. The rate of the variation of the parameters depends on the wavelength bandwidth of the light. The probability distribution governing the instantaneous elliptical polarization parameters of unpolarized light were worked out by Hurwitz (1944). He reached the conclusion that the median value for the ratio of the minor axis length to the major axis length is .268 . Thus for unpolarized light, more than half the time, the major axis is more than 3.5 times as long as the minor axis.

Jones (1941a) gives the following relations between the Jones vector components and the elliptical polarization parameters. Let $\delta$ be the difference between the x and y phases, $\delta=\operatorname{Phase}\left(\mathrm{E}_{\mathrm{y}}\right)-\operatorname{Phase}\left(\mathrm{E}_{\mathrm{x}}\right)$. Let $\tan \theta$ be the ratio of the axes of the ellipse, and let $\psi$ be the orientation of the major axis, measured counterclockwise from the x axis. Then,

$$
\tan 2 \psi=\tan 2 \alpha \cos \delta,
$$

and,

$$
\cos 2 \theta=\sin 2 \alpha|\sin \delta|
$$

where,

$$
\tan \alpha=\frac{E_{y}}{E_{x}} .
$$

## Basis Jones Vectors

Table 4 lists the Jones vectors for the most common polarization states: horizontal linear, vertical linear, $+45^{\circ}$ linear, $-45^{5}$ linear, right circuilar ana ieft circular polarized light. These vectors can be multiplied by an arbitrary phase factor, $\mathrm{e}^{\mathrm{j} \delta}$. without changing the polarization form of the light; it only changes the absolute phase.

## Generalizations of the Jones Vector

Jones (1942) has derived extensions to the algebra of Jones vectors which allow it to handle the incoherent addition of light fields, such as unpolarized light, where $\epsilon_{\mathrm{x}}(\mathrm{t})$ and $\epsilon_{\mathrm{y}}(\mathrm{t})$ have variations on the order of a radian or greater during an optical period, $t=1 / \omega$. These problems are often handled with Stokes vectors and Mueller matrices, but since a relationship exists between the two formalisms. incoherent light can be treated with either calculus.

Jones (1942) also treats the case of changing the basis of the Jones calculus so that the basis states are different, ie. $(1,0) \neq$ horizontal linear polarized light, and $(0,1)$ \# vertical linearly polarized light, wut instead represent any orthogonal pair of polarization states, such as left and right circularly polarized light. The conclusion of Jones" study of this change of basis is that it adds no new physics or insight but does mathematically simplify certain problems.

## TABLE 4

## BASIS JONES VECTORS

## Linear Polarized Light

$$
\begin{array}{cc}
\text { Horizontal } & \text { Vertical } \\
\hat{\mathrm{H}}=\left[\begin{array}{l}
1 \\
0
\end{array}\right] & \hat{\mathrm{V}}=\left[\begin{array}{l}
0 \\
1
\end{array}\right] \\
+45 \text { Degrees } & -45 \text { Degrees } \\
\hat{\jmath}=\frac{\sqrt{2}}{2}\left[\begin{array}{l}
1 \\
1
\end{array}\right] & \hat{\imath}=\frac{\sqrt{2}}{2}\left[\begin{array}{r}
1 \\
-1
\end{array}\right]
\end{array}
$$

## Circularly Polarized Light

$$
\begin{array}{cc}
\text { Right } & \text { Left } \\
\hat{\mathrm{R}}=\frac{\sqrt{2}}{2}\left[\begin{array}{c}
1 \\
-\mathrm{j}
\end{array}\right] & \hat{\mathrm{L}}=\frac{\sqrt{2}}{2}\left[\begin{array}{l}
1 \\
\mathrm{j}
\end{array}\right]
\end{array}
$$

## The Jones Matrix and the C Vector

## Definition of the Jones Matrix in terms of the Jones Vector

In his original paper, Jones (1941a) shows that the relationship between the Jones vector incident on a polarizer, $\mathbf{J}$, and the Jones matrix transmitted or reflected by a polarizer, $\vec{J}^{\prime}$, can always be related by a matrix, the Jones matrix, J. Only certain transformations of the field components are allowed, those describable by a mairix. The fundamenial reiaionsnip beiween the vecior componenis of the electromagnetic fields before and after a polarizing element is.

$$
\overrightarrow{\mathbf{J}}=\mathbf{J} \mathbf{J} .
$$

The Jones matrix, $\mathbf{J}$, is a two by two matrix with complex elements,

$$
J=\left[\begin{array}{ll}
j_{11} & j_{12} \\
j_{21} & j_{22}
\end{array}\right], j_{k, 1}=a_{k, 1}+j_{k_{k}, 1} .
$$

Thus the Jones matrix has eight degrees of freedom. Thus there are eight different forms of polarization behavior, a concept that will be developed further. Every Jones matrix corresponds to a physically realizable polarizer.

## Cascaded Polarizers

The Jones matrix associated with an optical ray path through a sequence of polarizers is just the matrix product of the Jones matrices for the individual polarizers. If an optical ray traverses a series of elements, $1,2, \ldots \mathbf{Q}$, and the Jones matrices appropriate to that ray for each element are, $\mathbf{J}_{1}, \mathbf{J}_{\mathbf{2}}, \ldots \mathbf{J}_{\mathbf{Q}}$, then the Jones matrix describing the polarization properties of the system along this ray path is given by the matrix product.

$$
\mathbf{J}=\mathbf{J}_{Q} \ldots \mathbf{J}_{2} \mathbf{J}_{1}=\prod_{q=Q,-1}^{1} \mathbf{J}_{\mathbf{q}}
$$

Since the Jones matrix of an optical element is dependent upon the wavelength, angle of incidence, orientation, and path through the element, care must be exercised in using the correct Jones matrix for a given optical ray. Only for a collimated monochromatic beam at normai incidence tinrougii a series of planar optical interfaces can a single Jones matrix can be written for the entire cross section of the beam. For nonplanar surfaces or nonplanar wavefronts or multiple wavelengths, different parts of the beams have different Jones matrices describing their interaction with the polarization elements.

## Coordinate System

The coordinate system of the Jones matrix is defined in terms of the $x-y$ coordinates for the Jones vector.

It is often desirable to align the Jones calculus coordinates with the $s$ and $p$ planes of an optical interface, since most thin film polarization equations are defined with respect to the $s$ and $p$ planes. Only for plane surfaces does the orientation of the $s$ and $\overline{\boldsymbol{p}}$ plazes remain fixed across the surface. For nonplanar surfaces, it is necessary to maintain two sets of coordinates, the global $x$ and $y$ coordinates with respect to which the Jones matrix is defined, and a local $s$ and $p$ coordinate aat each point on the interface. The local $s$ and $p$ coordinate system have the $x^{\prime}$ and $y^{\prime}$ axes aligned with the local $s$ and $p$ planes of the surface. The Jones matrix for an interface can be evaluated in the $\mathbf{x}^{\prime}-\mathbf{y}^{\prime}$ coordinates then this s-p Jones matrices can be rotated to bring it into the global $x-y$ coordinate system.

## Basis States

The basis states of the Jones matrix can be considered as the matrices:

$$
\begin{array}{ll}
\mathbf{B}_{1}=\left[\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right], & \mathbf{B}_{2}=\left[\begin{array}{ll}
0 & 1 \\
0 & 0
\end{array}\right] . \\
\mathbf{B}_{3}=\left[\begin{array}{ll}
0 & 0 \\
1 & 0
\end{array}\right], & \mathbf{B}_{4}=\left[\begin{array}{ll}
0 & 0 \\
0 & 1
\end{array}\right] . \\
\mathbf{B}_{5}=\left[\begin{array}{ll}
\mathbf{j} & 0 \\
0 & 0
\end{array}\right], & \mathbf{B}_{6}=\left[\begin{array}{ll}
0 & \mathbf{j} \\
0 & 0
\end{array}\right] . \\
\mathbf{B}_{7}=\left[\begin{array}{ll}
0 & 0 \\
\mathbf{j} & 0
\end{array}\right], & \mathbf{B}_{8}=\left[\begin{array}{ll}
0 & 0 \\
0 & j
\end{array}\right] .
\end{array}
$$

Then an arbitrary Jones matrix can be expressed as,

$$
J=\sum_{k=1}^{8} b_{k} B_{k}
$$

The basis states correspond to the following combinations of polarizers:
(1) A linear polarizer aligned to transmit along the x axis,
(2) A half wave plate oriented at 45 degrees followed by a linear polarizer transmitting along the $y$ axis.
(3)
(4)

A half wave plate oriented at 45 degrees followed by a linear polarizer transmitting along the x axis.

A linear polarizer aligned to transmit along the $y$ axis.
(5) to (8) The same as (1) to (4) eycept that all polarizers are followed by a quarter wave thickness of nonpolarizing material (nonbirefringent, nondichroic.)

The B's comprise a convenient basis for the Jones calculus but are not the most convenient basis for understanding polarizer problems. In particular, the B's are all singular matrices with zero determinant and no matrix inverse.

## Pauli Spin Matrix Basis and the C Vector

The Pauli spin matrices form a much more useful basis for the Jones matrix "space." The Pauli spin matrix basis was originally introduced by Jones (1948) with his "N-matrices." The approach presented here continues in the spirit with which Jones first introduced this basis. The expansion of the Jones matrices with the Pauli spin matrix basis set is especially concise for the theoretical development of the polarization aberrations.

The identity matrix, $\sigma_{0}$ and the Pauli spin matrices, $\sigma_{1}, \sigma_{2}$, and $\sigma_{3}$, are defined in Table 5.

For the characterization of polarizers, it is appropriate to label the spin matrices differently than is customary in quantum mechanics, since the labels $\mathrm{x}, \mathrm{y}$ and $z$, as used in quantum mechanics, are misleading here; for example: CohenTannoudji, Diu and Laloe (1977, page 417). Gottfried (1966, page 275), Landau and Lifshitz (1977, page 202). The foliowing expressions relate the present definitions to the quantum mechanics convention:

$$
\sigma_{1}=\sigma_{\mathrm{z}}, \sigma_{2}=\sigma_{\mathrm{x}}, \sigma_{\mathrm{s}}=\sigma_{\mathrm{y}} .
$$

Using the Pauli spin matrix basis, an arbitrary Jones matrix is expressed as

$$
\mathrm{J}=\sum_{\mathrm{k}=0}^{3} \mathrm{c}_{\mathrm{k}} \sigma_{\mathrm{k}}
$$

The c's can be gathered into a four element vector with complex elements called the " C vector." The C vector expression

$$
\vec{C}=\left[c_{0}, c_{1}, c_{2}, c_{3}\right] .
$$

represents the Jones matrix,

## TABLE 5

THE IDENTITY MATRIX AND THE PAULI SPIN MATRICES

$$
\begin{aligned}
& \sigma_{0}=\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] \\
& \sigma_{1}=\left[\begin{array}{rr}
1 & 0 \\
0 & -1
\end{array}\right] \\
& \sigma_{2}=\left[\begin{array}{ll}
0 & 1 \\
1 & 0
\end{array}\right] \\
& \sigma_{3}=\left[\begin{array}{rr}
0 & -j \\
j & 0
\end{array}\right]
\end{aligned}
$$

$$
J=c_{0} \sigma_{0}+c_{1} \sigma_{1}+c_{2} \sigma_{2}+c_{3} \sigma_{3}=\left[\begin{array}{cc}
c_{0}+c_{1} & c_{2}-j c_{3} \\
c_{2}+j c_{3} & c_{0}-c_{1}
\end{array}\right] .
$$

When needed, $\rho_{\mathrm{i}}$ and $\phi_{\mathrm{i}}$ will refer to the amplitude and phase portions of the C vector elements,

$$
\vec{C}=\left[\rho_{0} \mathrm{e}^{\mathrm{j} \phi_{0}}, \rho_{1} \mathrm{e}^{\mathrm{j} \phi_{1}}, \rho_{2} \mathrm{e}^{\mathrm{j} \phi_{2}}, \rho_{3} \mathrm{e}^{\mathrm{j} \phi_{3}}\right]
$$

The elements of $\vec{C}$ are related to the Jones matrix elements by the equations:

$$
\begin{aligned}
& c_{0}=\frac{j_{11}+j_{22}}{2} \quad, c_{1}=\frac{j_{11}-j_{22}}{2} \\
& c_{2}=\frac{j_{12}+j_{21}}{2} \quad, c_{3}=\frac{j_{12}-j_{21}}{-2 j}
\end{aligned}
$$

The elements of the Jones matrix are related to the elements of $\vec{C}$ by the equations:

$$
\begin{array}{ll}
j_{11}=c_{0}+c_{1} & , j_{12}=c_{2}-j c_{i} . \\
j_{21}=c_{2}+j c_{3} & , j_{22}=c_{0}-c_{1} .
\end{array}
$$

The $C$ vector, like the Jones matrix, has eight degrees of freedom. The $C$ vector is used as a shorthand for the expansion of a Jones matrix into Pauli spin matrices. A vector equivalent to the $\mathbf{C}$ vector has been introduced in quantum mechanics for a similar purpose; see the section, "A convenient basis for the $2 \times 2$ matrix space" in Quantum Mechanics, Cohen-Tannoudji, Diu and Laloe (1977, pg. 419).

## The Jones Matrix and C Vectors for Specific Polarizers

Every Jones matrix represents a physically realizable polarizer. Tables of Jones matrices for various polarizers are found in Azzam and Bashara (1977. Section 2.2.3), Hecht and Zajac (1974, Table 8.6), Shurcliff (1962, Appendix 2), and

Theocaris and Gdoutos (1979, Table 4.1). Table 6 is a listing of the Jones matrices and C vectors for the most common polarizers and retarders.

## Normalized Form for the C Vectors

A very useful normalized form for the $\mathbf{C}$ vectors factors $\mathrm{c}_{\mathrm{o}}$ outside the brackets and treated it as a constant. This constant is labeled $\tau$ to avoid confusion with the notimalized first eiement inside ine brackets. The normailization constant, $\tau$, is a complex number which is the amplitude and phase transmission of the optical element in the absence of polarization. If

$$
\vec{C}=\left[c_{0}{ }^{\prime}, c_{1}^{\prime}, c_{2}^{\prime}, c_{3}^{\prime}\right]
$$

is an unnormalized $C$ vector, theu the equivalent normalized $C$ vector is

$$
\overrightarrow{\mathbf{C}}=\tau\left[1, c_{1}, c_{2}, c_{3}\right] .
$$

with

$$
\tau=c_{0}^{\prime}, c_{1}=\frac{c_{1}^{\prime}}{\tau}, c_{2}=\frac{c_{2}^{\prime}}{\tau}, \text { and, } c_{3}=\frac{c_{3}^{\prime}}{\tau} .
$$

A C vector with $\mathrm{c}_{0}=0$ obviously cannot be normalized in this fashion.

## Pauli Spin Matrices

The following identities and properties of the Pauli spin matrices are collected here for reference.

Let $\alpha, \beta, \gamma \in\{1,2,3\}$, the indices for the three Pauli spin matrices.

## The Ideutity Matrix

The basis matrix $\sigma_{0}$ is the identity element. Let $M$ be an arbitrary matrix; then

$$
\sigma_{0} \mathbf{M}=\mathbf{M} \sigma_{0}=\mathbf{M} .
$$

## TABLE 6

## JONES MATRICES AND C VECTORS FOR IDEAL POLARIZERS

## Linear Polarizers

Angle of Transmission
Axis
Jones Matrix
C Vector
$0^{0}$

$$
\left[\begin{array}{ll}
1 & 0 \\
0 & 0
\end{array}\right] \quad \frac{1}{2}[1,1,0,0]
$$

$45^{\circ}$
$\frac{1}{2}\left[\begin{array}{ll}1 & 1 \\ 1 & 1\end{array}\right] \quad \frac{1}{2}[1,0,1,0]$
$90^{\circ}$
$\left[\begin{array}{ll}0 & 0 \\ 0 & 1\end{array}\right] \quad \frac{1}{2}[1,-1,0,0]$
$135^{\circ}$
$\frac{1}{2}\left[\begin{array}{rr}1 & -1 \\ -1 & 1\end{array}\right] \quad \frac{1}{2}[1,0,-1,0]$

Circular Polarizers

$$
\text { Jones Matrix } \quad \text { C Vector }
$$

Left
$\frac{1}{2}\left[\begin{array}{rr}1 & -j \\ j & 1\end{array}\right] \quad \frac{1}{2}[1,0,0,1]$

Right

$$
\frac{1}{2}\left[\begin{array}{rr}
1 & j \\
-j & 1
\end{array}\right] \quad \frac{1}{2}[1,0,0,-1]
$$

TABLE 6-Continued

JONES MATRICES AND C VECTORS FOR IDEAL POLARIZERS

## Retarders

## Quarter Wave Linear Retarders

Fast Axis
Angle

|  | Jones Matrix | C Vector |
| :--- | :--- | :--- |
| $0^{\circ}$ | $\frac{\sqrt{2}}{2}\left[\begin{array}{cc}1+\mathrm{j} & 0 \\ 0 & 1-\mathrm{j}\end{array}\right]$ | $\frac{\sqrt{2}}{2}[1, \mathrm{j}, 0,0]$ |
| $45^{\circ}$ | $\frac{\sqrt{2}}{2}\left[\begin{array}{cc}1 & \mathrm{j} \\ \mathrm{j} & 1\end{array}\right]$ | $\frac{\sqrt{2}}{2}[1,0, \mathrm{j}, 0]$ |
| $90^{\circ}$ | $\frac{\sqrt{2}}{2}\left[\begin{array}{rr}1-\mathrm{j} & 0 \\ 0 & 1+\mathrm{j}\end{array}\right]$ | $\frac{\sqrt{2}}{2}[1,-\mathrm{j}, 0,0]$ |
| $135^{\circ}$ | $\frac{\sqrt{2}}{2}\left[\begin{array}{cc}1 & -\mathrm{j} \\ -\mathrm{j} & 1\end{array}\right]$ | $\frac{\sqrt{2}}{2}[1,0,-\mathrm{j}, 0]$ |

## Quarter Wave Circular Retarders

Jones Matrix $\quad \underline{C}$ Vector
Left $\quad \frac{\sqrt{2}}{2}\left[\begin{array}{rr}1 & 1 \\ -1 & 1\end{array}\right] \quad \frac{\sqrt{2}}{2}[1,0.0, \mathrm{j}]$
Right $\quad \frac{\sqrt{2}}{2}\left[\begin{array}{cc}1 & -1 \\ 1 & 1\end{array}\right] \quad \frac{\sqrt{2}}{2}[1,0,0,-j]$

TABLE 6-Continued

JONES MATRICES AND C VECTORS FOR IDEAL POLARIZERS

## Half Wave Linear Retarciers

Fast Axis
Angle

$$
\text { Jones Matrix } \quad \text { C Vector }
$$

$0^{\circ}$ or $90^{\circ}$
$\left[\begin{array}{rr}\mathbf{j} & \mathbf{0} \\ 0 & -\mathbf{j}\end{array}\right]$
$[0, j, \hat{u}, 0]$
$\pm 45^{\circ}$

$$
\left[\begin{array}{ll}
0 & \mathrm{j} \\
\mathrm{j} & 0
\end{array}\right] \quad[0,0, \mathrm{j}, 0]
$$

Half Wave Circular Retarders
Jones Matrix C Vector
Left or Right

$$
\left[\begin{array}{rr}
0 & 1 \\
-1 & 0
\end{array}\right] \quad[0,0,0, j]
$$

## Trace and Determinant

$$
\begin{gathered}
\text { Trace } \sigma_{\alpha}=0=\sigma_{11}+\sigma_{22} \\
\text { Det } \sigma_{\alpha}=-1=\sigma_{11} \sigma_{22}-\sigma_{12} \sigma_{21}
\end{gathered}
$$

## Anticommutativity

$$
\sigma_{\alpha} \sigma_{\beta}=-\sigma_{\beta} \sigma_{\alpha} .
$$

## Relation to the Identity Matrix

$$
\sigma_{\alpha} \sigma_{\alpha}=\sigma_{0}
$$

From the last two relations, the following is derived,

$$
\sigma_{\alpha} \sigma_{\beta} \sigma_{\gamma}=j \sigma_{0}
$$

In addition, when ( $\alpha, \beta, \gamma$ ) is an even permuitation of ( $1,2,3$ ), then

$$
\sigma_{\alpha} \sigma_{\beta}=j \sigma_{\gamma}
$$

This produces the relations:

$$
\begin{aligned}
& \sigma_{1} \sigma_{2}=\mathrm{j} \sigma_{3}, \sigma_{2} \sigma_{1}=-\mathrm{j} \sigma_{3}, \\
& \sigma_{2} \sigma_{3}=\mathrm{j} \sigma_{1}, \sigma_{3} \sigma_{2}=-\mathrm{j} \sigma_{1}, \\
& \sigma_{3} \sigma_{1}=\mathrm{j} \sigma_{2}, \sigma_{1} \sigma_{3}=-\mathrm{j} \sigma_{2} .
\end{aligned}
$$

Chapter 4 explores these relations, which govern the order dependent properties of polarizers and retarders, in detail.

## Eigenvalues and Eigenvectors

Table 7 contains the eigenvalues and eigenvectors of the Pauli spin matrices in terms of the basis Jones vectors defined in Table 4.

Any vector is an eigenvector of $\sigma_{0}$, the identity matrix, with an eigenvalue equal to one.

TABLE 7

EIGENVALUES AND EIGENVECTORS OF THE PAULI SPIN MATRICES

| Polarizer | Eigen- <br> value | Eigen- <br> vector | Eigen- <br> value | Eigen- <br> vector |
| :---: | :---: | :---: | :---: | :---: |
| $\sigma_{1}$ | 1 | $\hat{\mathbf{V}}$ | -1 | $\hat{\mathrm{H}}$ |
| $\sigma_{2}$ | 1 | $\hat{j}$ | -1 | $\hat{\mathrm{~L}}$ |
| $\sigma_{3}$ | 1 | $\hat{\mathrm{~L}}$ | -1 | $\hat{\mathbf{R}}$ |

## Effect of the Pauli Matrices on Basis Polarization States

Table 8 tabuiates the effect of the Pauli spin matrices as polarizers operating on the basis polarization states. The spin matrix is listed along the left side of the table and the incident polarization state along the top. The entries in the table are the transmitted polarization states.

This table illustrates the usefulness of the Pauli spin matrices as basis states for the Jones matrix space. When one of the basis polarization states is acted on by one of the Pauli basis polarizers, a mixing of states occurs only between orthogonal states. In the table the first two columns contain only $\hat{H}$ 's and $\hat{V}$ 's, the third and fourth columns, only $\hat{\mathrm{V}}$ 's and $\hat{\jmath} \mathrm{f}$, and the last two columns, only $\hat{\mathrm{L}}$ 's and $\hat{\mathrm{R}}$ 's. When passing $\hat{H}$ through one of the basis polarizers, it never couples directly into $\hat{R}, \hat{L}, \hat{\jmath}$ or $\hat{\mathrm{V}}$; it only couples into $\hat{\mathrm{H}}$ or $\hat{\mathrm{V}}$. This is a substantial simplification.

## The Exponential of a Matrix

The exponential of the matrix $M$ times a constant $\alpha$ is defined as

$$
\begin{aligned}
\exp (\alpha \mathbf{M})=\sigma_{0} & +\alpha \mathbf{M}+\frac{\alpha^{2} M^{2}}{2!}+\frac{\alpha^{3} M^{3}}{3!}+\ldots \\
& =\sum_{n=0}^{\infty} \frac{\frac{\mathfrak{x}^{n} M^{n}}{n!}}{n}
\end{aligned}
$$

where $\sigma_{0}$ is the identity matrix.

## TABLE 8

## Polarization States Transmitted by the Basis Polarizers



## Exponentials of the Basis Matrices

The complex exponentials of the Pauli spin matrices are especially useful, since the Pauli spin matrices squared equal the identity matrix:

$$
\begin{array}{ll}
\exp (\alpha \sigma) & =e^{\alpha} \sigma_{0}, \\
\exp \left(\mathrm{j} \alpha \sigma_{0}\right) & =\sigma_{0} \cos \alpha+\mathrm{j} \sigma_{0} \sin \alpha, \\
\exp \left(\alpha \sigma_{1}\right) & =\sigma_{0}+\alpha \sigma_{1}+\frac{\alpha^{2} \sigma_{0}}{2}+\frac{\mathrm{j} \alpha^{3} \sigma_{1}}{3!}+\ldots \\
& =\sigma_{0} \cosh \alpha+\sigma_{1} \sinh \alpha, \\
\exp \left(\mathrm{j} \alpha \sigma_{1}\right) & =\sigma_{0}+j \alpha \sigma_{1}-\frac{\alpha^{2} \sigma_{0}}{2}-\frac{\mathrm{j} \alpha^{3} \sigma_{1}}{3!}+\ldots \\
& =\sigma_{0} \cos \alpha+j \sigma_{1} \sin \alpha, \\
\exp \left(\alpha \sigma_{2}\right) & =\sigma_{0} \cosh \alpha+\sigma_{2} \sinh \alpha, \\
\exp \left(j \alpha \sigma_{2}\right) & =\sigma_{0} \cos \alpha+j \sigma_{2} \sin \alpha, \\
\exp \left(\alpha \sigma_{3}\right) & =\sigma_{0} \cosh \alpha+\sigma_{3} \sinh \alpha, \\
\exp \left(j \alpha \sigma_{3}\right) & =\sigma_{0} \cos \alpha+j \sigma_{3} \sin \alpha,
\end{array}
$$

## Properties of the Jones Matrix and the C Vector

## Rotated Polarizers

If a polarizer with Jones matrix $\mathbf{J}$ is rotated through an angle $\theta$ (positive if counterclockwise), the Jones matrix becomes

$$
\mathbf{J}^{\prime}(\theta)=\mathbf{R}(\theta) \mathbf{J} \mathbf{R}(-\theta) .
$$

The $\mathbf{R}(\theta)$ 's are the Jones rotation matrices:

$$
\mathbf{R}(\theta)=\left[\begin{array}{rr}
\cos \theta & -\sin \theta \\
\sin \theta & \cos \theta
\end{array}\right] \text {, and, } \mathbf{R}(-\theta)=\left[\begin{array}{rr}
\cos \theta & \sin \theta \\
-\sin \theta & \cos \theta
\end{array}\right] .
$$

The Jones rotation matrices obey the relations,

$$
\mathbf{R}(\alpha) \mathbf{R}(\beta)=\mathbf{R}(\beta) \mathbf{R}(\alpha)=\mathbf{R}(\alpha+\beta) \text { and, } \mathbf{R}(\alpha) \mathbf{R}(-\alpha)=\sigma_{0} .
$$

## Rotation of the Basis Matrices

The identity matrix is invariant under rotation;

$$
\mathbf{R}(\theta) \sigma_{0} \mathbf{R}(-\theta)=\sigma_{0} .
$$

Under rotation, $\sigma_{1}$ and $\sigma_{2}$ couple into each other;

$$
\begin{aligned}
& \mathbf{R}(\theta) \sigma_{1} \mathbf{R}(-\theta)=\sigma_{1} \cos 2 \theta+\sigma_{2} \sin 2 \theta, \\
& \mathbf{R}(\theta) \sigma_{2} \mathbf{R}(-\theta)=-\sigma_{1} \sin 2 \theta+\sigma_{2} \cos 2 \theta .
\end{aligned}
$$

$\sigma_{3}$ is invariant under rotation;

$$
\mathbf{R}(\theta) \sigma_{3} \mathbf{R}(-\theta)=\sigma_{3} .
$$

## $\underline{\text { Rotation of a C Vector }}$

The $C$ vector for a rotated polarizer can be found from the rotated Jones matrix

$$
\begin{aligned}
\mathbf{R}(\theta) \mathbf{J} \mathbf{R}(-\theta) \quad & =\mathbf{R}(\theta)\left[\sum_{k=1}^{4} c_{k} \sigma_{k}\right] \mathbf{R}(-\theta) \\
& =c_{0} \sigma_{0}+\left(c_{1} \cos 2 \theta-c_{2} \sin 2 \theta\right) \sigma_{1}+\left(c_{1} \sin 2 \theta+c_{2} \cos 2 \theta\right) \sigma_{2}+c_{3} \sigma_{3}
\end{aligned}
$$

Thus, the C vector for a rotated polarizer is

$$
\vec{C}^{\prime}=\left[c_{0}{ }^{\prime}, c_{1}{ }^{\prime}, c_{2}{ }^{\prime}, c_{3}^{\prime}\right]=\left[c_{0}, c_{1} \cos 2 \theta-c_{2} \sin 2 \theta, c_{1} \sin 2 \theta+c_{2} \cos 2 \theta, c_{3}\right] .
$$

The rotated C vector is given by the following product,

$$
\vec{C}^{\prime}=\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
0 & \cos 2 \theta & -\sin 2 \theta & 0 \\
0 & \sin 2 \theta & \cos 2 \theta & 0 \\
0 & 0 & 0 & 1
\end{array}\right]\left[\begin{array}{l}
c_{0} \\
c_{1} \\
c_{2} \\
c_{3}
\end{array}\right] .
$$

The matrix $\mathbf{R}_{\mathrm{C}}(\theta)$ is the rotation operator for $\mathbf{C}$ vectors,

$$
\mathbf{R}_{\mathbf{c}}(\theta)=\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
0 & \cos 2 \theta & -\sin 2 \theta & 0 \\
0 & \sin 2 \theta & \cos 2 \theta & 0 \\
0 & 0 & 0 & 1
\end{array}\right] .
$$

## The C Vector for Cascaded Polarizers

The C vector for the matrix product of two C vectors is obtained from the application of the Pauli spin matrix identities. Let

$$
\vec{C}=\left[c_{0}, c_{1}, c_{2}, c_{3}\right] \text { and, } \vec{C}^{\prime}=\left[c_{0}^{\prime}, c_{1}^{\prime}, c_{2}^{\prime}, c_{3}^{\prime}\right] .
$$

The $C$ vector $\vec{C}^{\prime \prime}$ which corresponds to light operated on by $\vec{C}$ ', then $\vec{C}$, has the Jones matrix,

$$
J^{\prime \prime}=J^{\prime} J^{\prime}=\left[\sum_{i=1}^{4} \mathrm{c}_{\mathrm{i}} \sigma_{\mathrm{i}}\right]\left[\sum_{\mathrm{i}=1}^{4} \mathrm{c}_{\mathrm{i}}^{\prime} \sigma_{\mathrm{i}}\right]=\sum_{\mathrm{i}=1}^{4} \mathrm{c}_{\mathrm{i}} \boldsymbol{\sigma}_{\mathrm{i}} .
$$

By equating terms in the previous equation, $\vec{C}$ " can be found from the following matrix vector product,

$$
\begin{aligned}
\overrightarrow{\mathbf{~}}^{\prime \prime} & =\left[\begin{array}{l}
\mathrm{c}^{\prime \prime}{ }_{0} \\
\mathrm{c}_{1}{ }_{1} \\
\mathrm{c}_{2} \\
\mathrm{c}_{2}{ }_{3}
\end{array}\right]=\mathbf{K}(\overrightarrow{\mathrm{C}})\left[\begin{array}{l}
\mathrm{c}_{0}^{\prime} \\
\mathrm{c}_{1}^{\prime} \\
\mathrm{c}_{1}^{\prime} \\
\mathrm{c}_{3}^{\prime}
\end{array}\right] \\
& =\left[\begin{array}{rrrr}
\mathrm{c}_{0} & \mathrm{c}_{1} & \mathrm{c}_{2} & \mathrm{c}_{3} \\
\mathrm{c}_{1} & \mathrm{c}_{0} & -\mathrm{jc}_{3} & \mathrm{jc}_{2} \\
\mathrm{c}_{2} & \mathrm{jc}_{3} & \mathrm{c}_{0} & -\mathrm{jc}_{1} \\
\mathrm{c}_{3} & -\mathrm{jc} & \mathrm{jc}_{1} & \mathrm{c}_{0}
\end{array}\right]\left[\begin{array}{l}
\mathrm{c}_{0}^{\prime} \\
\mathrm{c}_{1}^{\prime}{ }_{1} \\
\mathrm{c}_{2}^{\prime} \\
\mathrm{c}_{3}^{\prime}
\end{array}\right] .
\end{aligned}
$$

The matrix, $\mathbf{K}(\vec{C})$, the "polarization coupling matrix", contains the couplings between the basis polarization states. These couplings arise from the Pauli spin matrix identities.

## The Meaning of the Coefficients of the C Vector

The C vector is introduced to simplify the representation of polarizers. Each of the elements of the C vector represents a specific type of polarizer behavior. Table 9 is a list of the polarization property associated with each $C$ vector elements.

The real parts of the C vector correspond to amplitude effects, absorption and polarization. The phase portion of the $\mathbf{C}$ vector represent phase effects, propagation and birefringence. The first element, $\mathrm{c}_{0}=\rho_{0} \mathrm{e}^{+\mathrm{j} \phi_{0}}$, is the coefficient of the identity matrix. Thus it must represent effects that are polarization state independent; these are amplitude and phase. The last element, $\mathrm{c}_{3}=\rho_{3} \mathrm{e}^{+\mathrm{j} \phi_{3}}$, multiplies the spin matrix $\sigma_{3}$ which is rotation invariant. Thus the $\mathrm{c}_{3}$ term represents the circular polarization effects; $\rho_{3}$ describes circular polarization or circular dichroism and $\phi_{3}$ describes circular retardance or circular birefringence. The remaining two elements, $\mathrm{c}_{1}$ and $\mathrm{c}_{2}$, represent linear polarization. Linear terms require two degrees of freedom: magnitude and orientation. Thus, $\rho_{1}$ and $\rho_{2}$ characterize linear polarization or linear dichroism, $\rho_{1}$ in the $0^{\circ}$ and $90^{\circ}$ directions, $\rho_{2}$ in the $\pm 45^{\circ}$ directions. Likewise, $\phi_{1}$ and $\phi_{2}$ characterize linear retardance or linear birefringence. Appendix C contains a detailed discussion of the properties of the elements of the $C$ vector.

TABLE 9

## Interpretation of the C Vector Elements

| Miatrix | Coefficient | Meaning |  |
| :--- | :--- | :--- | :--- |
|  |  |  |  |
| $\sigma_{0}$ | $\rho_{0}$ | Amplitude | Absorption <br> $\sigma_{0}$ |
| $\phi_{0}$ | Phase | Phase |  |
| $\sigma_{1}$ | $\rho_{1}$ | Amplitude | Linear Polarization along Axes |
| $\sigma_{i}$ | $\phi_{i}$ | Phase | Linear Retardance along Axes |
| $\sigma_{2}$ | $\rho_{2}$ | Amplitude | Linear Polarization. 45 |
| $\sigma_{2}$ | $\phi_{2}$ | Phase | Linear Retardance, 45 |
| $\sigma_{3}$ | $\rho_{3}$ | Amplitude | Circular Polarization |
| $\sigma_{3}$ | $\phi_{3}$ | Phase | Circular Retardance |

## CHAPTER 3

## WEAK POLARIZERS

The remainder of this dissertation is a study of the polarization characteristics of optical systems comprised of coated and uncoated lens and mirror elements. The next five chapters develop a method for approximating the polarization behavior of these systems with a set of functions, the polarization aberrations, which complement the wavefront aberrations of geometrical optics. The final chapter treats the exact caicuiation of the polarization periormance of opical systems as predicted by polarization theory by the method of polarization ray tracing. Both methods have as their specific goal the determination of the instrumental polarization function, polarization matrices associated with arbitrary ray paths through the system.

## Instrumental Polarization Versus

## Transmitted Light Polarization Calculations

Two types of polarization calculations can be performed: instrumental polarization and transmitted light polarization. The first is the calculation of the instrumental polarization associated with ray paths through an optical system, the Jones matrix for a given ray. The other type of calculation determines the state of polarization, such as a Jones vector, transmitted by the system along a given ray path for a specified input polarization state. This work deals with the instrumental polarization calculation since it is more fundamental. Once the instrumental polarization function for the system is known, the transmitted Jones vectors are readily determined for all input polarization states. The instrumentai polarization function is the "potential function" for the transmitted polarization function.

## Transparent Systems

The emphasis of this research is on systems which are weakly polarizing and highly transparent. This includes most lenses, cameras, telescopes, microscopes and other optical systems which do not dispiay large amounts of absorption or polarization. These "transparent systems" do not contain linear polarizers, retarders, diffraction gratings, optically active crystals or similar elements. The ideal Jones matrix for a ray through a transparent nonpolarizing system is

$$
\mathrm{J}_{\text {ideal }}=\mathrm{e}^{\mathrm{j} \delta}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]
$$

where $\delta$ is the optical path length for the ray in radians. This ray Jones matrix is the identity matrix, which signifies that the system has no absorption or polarization. Since this is the desired form of the Jones matrix for a large class of systems, the approach developed here obtains the instrumental polarization function as a Taylor series in the ray coordinates about $\mathbf{J}_{\text {ideal }}$. This approach is easily be modified for systems which are not highly transparent or which contain strong polarizers by performing the Taylor series about the Jones matrix for the ray down the optical axis.

## S-P Coordinates

It is necessary to maintain two separate coordinate systems to effectively analyze problems involving light at nonnormal incidence at curved optical interfaces: $\mathrm{x}-\mathrm{y}$ coordinates and $\mathrm{s}-\mathrm{p}$ coordinates. The $\mathrm{x}-\mathrm{y}$ coordinates are the global $\mathrm{x}, \mathrm{y}$ and z coordinate system used to describe the optical system, with the $z$ axis coinciding with the optical axis of radially symmetric systems.

Polarization calculations with angle of incidence dependent polarizers are usually performed in s-p coordinates. The s-f coordinates are based on the concept of the $s$ and $p$ planes. Consider light with unit wavevector $\hat{k}$ (normalized to one)
incident at a surface with normal $\hat{\mathrm{n}}$. The plane of incidence, or "p plane" is the plane which contains $\hat{k}$ and $\hat{n}$. The angle of incidence $i$, the angle between $\hat{k}$ and $\hat{n}$, is

$$
i=\arccos [\hat{k} \cdot \hat{n}] .
$$

The plane perpendicular to the plane of incidence which contains $\hat{\mathrm{k}}$ is the "s plane." The unit vectors $\hat{\mathrm{k}}, \hat{\mathrm{s}}$, and $\hat{\mathrm{p}}$ form an orthonormal basis for the s-p coordinate system. The $\hat{s}$ basis vector is given by the cross product

$$
\hat{s}=\hat{k} \times \hat{n} .
$$

The splane is then spanned by $\hat{s}$ and $\hat{k}$. The $\hat{\mathrm{p}}$ vector is found from the GramSchmidt orthonormalization equation

$$
\hat{\mathbf{p}}=\frac{\hat{\mathrm{n}}-\hat{k}[\hat{k} \cdot \hat{\mathrm{n}}]}{|\hat{\mathrm{n}}-\hat{k}[\hat{k} \cdot \hat{\mathbf{n}}]|}
$$

The p plane is spanned by $\hat{\mathrm{p}}$ and $\hat{\mathrm{k}}$. For light at nonnormal incidence at an interface, the polarization behavior of the interiace is analyzed by resolving the $\vec{E}(t)$ into its $s$ and $p$ components:

$$
\begin{aligned}
& \vec{E}_{s}(t)=\hat{s}[\hat{s} \cdot \vec{E}(t)] \\
& \vec{E}_{p}(t)=\hat{p}[\hat{p} \cdot \vec{E}(t)] .
\end{aligned}
$$

The normalized Jones vector in s-p coordinates is defined as

$$
J=\frac{1}{\vec{E}(t)}\left[\begin{array}{l}
\vec{E}_{s}(t) \\
\vec{E}_{p}(t)
\end{array}\right]
$$

Since $\hat{s} \cdot \hat{n}=0, \hat{s}$ is tangential to (lies on) the surface. Thus $\vec{E}_{s}(t)$ drives any electric currents parallel to the surface. Since

$$
\hat{\mathbf{p}} \cdot \hat{\mathbf{n}} \neq 0
$$

the $p$ component is not tangential to the surface but has a component along $\hat{n}$. This normal electric field component tends to push electrons into the surface and then pull
them out of the surface. This difference between the $\hat{s}$ and $\hat{\mathbf{p}}$ components is the basic reason for the polarization differences between the $s$ and $p$ components of light at interfaces. As a mnemonic, the $p$ component is sometimes called the plunge component since it "plunges" into the surface, while the $s$ component is the skip component since it generally has the higher reflectance and thus "skips" off the surface.

As the angle of incidence approaches zero (normal incidence,) $\hat{\mathrm{p}}$ becomes tangent to the surface. At normal incidence, both $\hat{s}$ and $\hat{p}$ are tangent to the surface and the difference between the $s$ plane and $p$ plane becomes indeterminant. At normal incidence, all expressions for $s$ and $p$ polarization become equal for isotropic interfaces.

Most frequently, the functional form of the interface polarization is given in s-p coordinates. Typically, the Jones matrix for a ray at an optical interface is calculated in the s-p coordinates, then rotated into $x-y$ coordinates. Once the Jones matrices for the ray at all suzfaces are rotated into $x-y$ coordinates, they can be multiplied to yield the instrumental polarization along that ray path in the system x-y coordinates.

## Normalized C Vector

Wherever possible a normalized form of the $\mathbf{C}$ vector in s-p coordinates is used where the normal incidence amplitude and phase, $c_{0}(0)$ is factored out. The elements of the normalized $C$ vector are written as d's to distinguish them from unnormalized C vectors. The C vector for an angle of incidence dependent polarizer in s-p coordinates is $\quad{ }_{\mathrm{O}}^{\mathrm{i}}$ $)=\left[\mathrm{c}_{0}(\mathrm{i}), \mathrm{c}_{1}(\mathrm{i}), c_{2}(\mathrm{i}), c_{3}(\mathrm{i})\right]$.

The amplitude and phase transmittance at normal incidence is $\tau=c_{0}(0)$. The normalized $C$ vector is defined as

$$
\not{Z}(\mathrm{i})=\tau\left[1, d_{1}(\mathrm{i}), d_{2}(\mathrm{i}), d_{3}(\mathrm{i})\right] .
$$

with $d_{k}(i)=c_{k}(i) / c_{0}(0)$.

## Instrumental Polarization

The Jones matrix for all optical elements varies as the angle of incidence changes. Further, this change always involves more than just a variation in the intensity and phase of the light; it also involves polarization and retardance. A fine optical element used in a transparent system does not display polarization effects at normal incidence; it may show some absorption, reflection loss or phase shift, but not polarization or retardance. Its normalized $C$ vector at normal incidence is

$$
\vec{C}(\mathrm{i}=0)=\left[\begin{array}{ccc}
c_{0}, 0,0,0
\end{array}\right]=\tau\left[\begin{array}{lll}
1,0,0,0
\end{array}\right] .
$$

As the angle of incidence varies, the Jones matrix must also vary to satisfy the conditions of electromagnetic theory that the tangential components of $\vec{E}(t)$ and $\vec{H}(t)$ are matched across the interface. Ai an arbitrary angle of incidence $i$, the $C$ vector is of the form

$$
\left.\vec{Z}_{\mathrm{i}}\right)=\left[\mathrm{c}_{0}(\mathrm{i}), \mathrm{c}_{1}(\mathrm{i}), c_{2}(\mathrm{i}), c_{3}(\mathrm{i})\right]=\left[\rho_{0}(\mathrm{i}) \mathrm{e}^{\mathrm{j} \phi_{0}(\mathrm{i})}, \rho_{1}(\mathrm{i}) \mathrm{e}^{\mathrm{j} \phi_{1}(\mathrm{i})}, \rho_{2}(\mathrm{i}) \mathrm{e}^{\mathrm{j} \phi_{2}(\mathrm{i})}, \rho_{3}(\mathrm{i}) \mathrm{e}^{\mathrm{j} \phi_{3}(\mathrm{i})}\right] .
$$

The functional dependences of the C vector coefficients are calculated from the Fresnel equations or thin film equations for the interface.

## Weak Polarizers

A weak polarizer is defined as a polarization element having a $C$ vector such that

$$
c_{0} \gg \sqrt{\left|c_{1}\right|^{2}+\left|c_{2}\right|^{2}+\left|c_{s}\right|^{2}} .
$$

The retardances of a weak polarizar are all small.

$$
\phi_{1}, \phi_{2}, \phi_{3} \ll 1 \text { radian. }
$$

Similarly, the polarizance of a weak polarizer is close to zero. Such a weak polarization element transmits light in a polarization state similar to the incident state with only weak coupling into other polarization states. The polarization behavior is dominated by transmission with only traces of polarization or retardance. Any polarization present is at the few percent level or less, such that any linearly polarized incident beam has a transmission coefficient which varies a few percent or less with orientation. Similarly, the retardation is a few degrees or less, far less than a quarter wave retarder with 90 degrees of retardation. Near normal incidence, metals in reflection and dielectric refracting interfaces are weak polarizers. Near normal incidence, antireflection coated lenses used in transmission and metals with reflection enhancing coatings are typically weak polarizers for wavelengths near the thin film design wavelength.

## Amplitude Transmission Relations

Throughout this section, the plane of incidence is aligned with the y axis.
The amplitude transmission equations for an interface are the equations which relate the amplitude and phase of the electric fields of the incident, reflected and refracted beams at an interface. The most general amplitude transmission equations for a nonscattering linear interface are:

$$
\begin{aligned}
\vec{E}_{s}^{\prime} & =a_{s s} \vec{E}_{s}+a_{p s} \vec{E}_{p} \\
\vec{E}_{p}^{\prime} & =a_{s p} \vec{E}_{s}+a_{p p} \vec{E}_{p}
\end{aligned}
$$

This equation is equivalent to the Jones matrix equation

$$
\left[\begin{array}{l}
\vec{E}_{\mathrm{s}}^{\prime} \\
\overrightarrow{\mathbf{E}}_{\mathrm{p}}^{\prime}
\end{array}\right]=\left[\begin{array}{ll}
\mathrm{a}_{\mathrm{ss}} & \mathrm{a}_{\mathrm{ps}} \\
\mathrm{a}_{\mathrm{sp}} & \mathrm{a}_{\mathrm{pp}}
\end{array}\right]\left[\begin{array}{l}
\vec{E}_{\mathrm{s}} \\
\vec{E}_{\mathrm{p}}
\end{array}\right]
$$

For interfaces whese eigenpolarizations are linear polarized light oriented parallel and perpendicular to the plane of incidence, the transfer of energy across the interface is separable into two uncoupled components of the form:

$$
\begin{aligned}
& \vec{E}_{S}^{\prime}-a_{S} \vec{E}_{S}=\rho_{S} e^{i \delta \delta_{S}} \vec{E}_{S} . \\
& \vec{E}_{p}^{\prime}=a_{p} \vec{E}_{p}=\rho_{p} p^{j \delta_{p}} \vec{E}_{p} .
\end{aligned}
$$

The amplitude transmission coefficients $\mathrm{a}_{\mathrm{S}}$ and $\mathrm{a}_{\mathrm{p}}$, or in polar coordinates, $\rho_{\mathrm{S}}, \phi_{\mathrm{S}}, \rho_{\mathrm{p}}$ and $\phi_{p}$, are determined by the Fresnel equations or thin film equations for the interface. These equations, where the $s$ and $p$ equations are separable, are "separable amplitude transmission relations." Only polarizers with linearly polarized light as the eigenpolarizations have the energy transfer equations in the separable form. For example, uncoated lenses, metal mirrors and thin film coatings on metals or dielectrics where the coatings are homogeneous, isotropic and locally planar, have eigenpolarizations of linear polarized light in the $s$ and $p$ planes. Thus the transmission of light by these interfaces is described by separable amplitude transmission relations.

Separable amplitude tianismission relations correspond to a diagonal Jones matrices in s-p coordinates. The Jones matrix and $C$ vector for an separable amplitude transmission interface in s-p coordinates are:

$$
\begin{gathered}
J(i)=\left[\begin{array}{cc}
a_{s}(i) & 0 \\
0 & a_{p}(i)
\end{array}\right] . \\
\text { さ }=1 / 2\left[a_{s}(i)+a_{p}(i), a_{s}(i)-a_{p}(i), 0,0\right] .
\end{gathered}
$$

## Taylor Series Representation of Weak Polarizers

In geometrical aberration theory, expressions for the optical path length of ray segments through the the optical system are obtained by performing a Taylor series expansion on Snells law, the law of reflection and the grating equation, to obtain expressions for the optical path length as a power series expansion in the ray coordinates. Thus Snells law,

$$
n \sin i=n^{\prime} \sin i^{\prime} .
$$

is rewritten for $\mathrm{i}^{\prime}$ as

$$
i^{\prime}=\arcsin \left[\frac{n}{n^{\prime}} \sin i\right]=\frac{n}{n^{\prime}} i+\left[\frac{n^{3}}{n^{\prime 3}}-\frac{n}{n^{\prime}}\right] \frac{i^{3}}{6}+O\left\{i^{5}\right\} .
$$

The polarization aberrations are generated in an analogous fashion. To obtain the variation of the Jones matrix in the exit pupil of a system, the appropriate Fresnel equations or other polarization equations are required in Taylor series form. For radially symmetric optical systems, expansions in the angle of incidence about normal incidence are used.

The Taylor series of a function about zero is defined as

$$
\begin{aligned}
f(i) & \cong \sum_{\gamma=0}^{\infty}\left[\frac{\partial^{\gamma} f(i)}{\partial i^{\gamma}}\right]_{i=0} \frac{i^{\gamma}}{\gamma!} \\
& =f(0)+i\left[\frac{\partial f(i)}{\partial i}\right]_{i=0}+\frac{i^{2}}{2}\left[\frac{\partial^{2} f(i)}{\partial i^{2}}\right]_{i=0}+\ldots \\
& =f_{0}+f_{1} i+f_{2} i^{2}+\ldots=\sum_{\gamma=0}^{\infty} f_{\gamma^{i}} \gamma
\end{aligned}
$$

where the $f_{\gamma}$ are the Taylor series coefficients. The order of the Taylor series terms is given by the subscript $\gamma$.

An isotropic interface appears unchanged as it is rotated about the surface normal. Diffraction gratings, holograms and scratched surfaces are not isotropic. Some crystal surfaces and thin film coatings are not isotropic. For isotropic interfaces, the Fresnel equations are even functions since the surface, does not distinguish between angles of incidence of $+i$ and $-i$. Thus, $f(i)=f(-i)$, which is the definition of an even function.

An even function contains only even terms in its Taylor series expansion about the origin. Thus, the Taylor series representations of the Fresnel equations has the form,

$$
\begin{aligned}
f(i) \quad & \cong \sum_{\gamma=0,2}^{\infty}\left[\frac{\partial^{\gamma} f(i)}{\partial i^{\gamma}}\right]_{i=0} \frac{i^{\gamma}}{\gamma!} \\
& =f(0)+\frac{i^{2}}{2}\left[\frac{\partial^{2} f(i)}{\partial i^{2}}\right]_{i=0}+\frac{i^{4}}{4!}\left[\frac{\partial^{4} f(i)}{\partial i^{4}}\right]_{i=0}+\ldots \\
& =f_{0}+f_{2} i^{2}+f_{4} i^{4}+\ldots=\sum_{\gamma=, 20}^{\infty} f_{\gamma^{1}}{ }^{\gamma} .
\end{aligned}
$$

For weakly polarizing interfaces described by amplitude transmittance relations, the Taylor series forms of the Jones matrix and $C$ vector are calculated as follows. First, the Taylor series is determined for the amplitude transmission relations:

$$
a_{S}(i) \cong a_{S, 0}+a_{s, 2} i^{2}+a_{s, 4} i^{4}+\ldots \quad, \quad a_{p}(i) \cong a_{p, 0}+a_{p, 2} i^{2}+a_{p, 4} i^{4}+\ldots
$$

Then, the Taylor series expansion about $\mathrm{i}=0$ in s-p coordinates for the Jones matrix is

$$
J(i)=\left[\begin{array}{cc}
a_{s, 0} & 0 \\
0 & a_{p, 0}
\end{array}\right]+i^{2}\left[\begin{array}{cc}
a_{s, 2} & 0 \\
0 & a_{p, 2}
\end{array}\right]+\ldots .
$$

The corresponding $\mathbf{C}$ vector expansion in s-p coordinates is

$$
\vec{C}=\left[c_{00}+c_{02} i^{2}+\ldots, c_{10}+c_{12} i^{2}+\ldots, 0,0\right] .
$$

where:

$$
c_{0, \gamma}=\frac{1}{2}\left(a_{s, \gamma}+a_{3, \gamma}\right), c_{1, \gamma}=\frac{1}{2}\left(a_{s, \gamma}-a_{s, \gamma}\right) .
$$

For an interface characterized by separable amplitude transmission relations, the diagonal and circular polarization components, $c_{2}(i)$ and $c_{3}(i)$, are always zero.

The normalized $\mathbf{C}$ vector for the separable amplitude transmission relations is

$$
\overrightarrow{\mathrm{C}}=\tau\left[1+\mathrm{d}_{02} \mathrm{i}^{2}+\ldots, \mathrm{d}_{10}+\mathrm{d}_{12} \mathrm{i}^{2}+\ldots, 0,0\right],
$$

where

$$
\tau=c_{00}, \text { and, } d_{k, n}=\frac{c_{k, \gamma}}{c_{00}}
$$

The Jones matrix and $\mathbf{C}$ vector for coordinates other than the s-p coordinates are obtained from the polarization rotation operation. For example, the s-p coordinates are rotated with respect to the $x-y$ coordinates by $\theta$, the orientation of the piane of incidence. The $x-y$ Jones matrix and $s-p$ Jones matrix are related by

$$
\mathbf{J}_{\mathrm{xy}}=\mathbf{R}(-\theta) \mathbf{J}_{\mathrm{sp}} \mathbf{R}(\theta)
$$

## Obtaining the Taylor Coefficients from Sampled Data

The direct method for calculating the coefficients of a Taylor series given in the last section is impractical for many interfaces due to the complexity of calculating the partial derivatives for the appropriate amplitude transmission equations such as muitilayer thin film coatings.

The Taylor series coefficients can be obtained numerically from the $s$ and $p$ amplitude transmissions evaluated at a series of angles ô incidence. The following algorithm is given to sixth order but can be extended to higher order.

Let the calculated values of a function $F(x)$ be given at: $0, x, 2 x$, and $3 x$. Assume that $\mathrm{F}(\mathrm{x})$ is accurately represented by a sixth order even Taylor series,

$$
F(x)=f_{0}+f_{z} x^{2}+f_{4} x^{4}+f_{6} x^{6}
$$

whose coefficients are to be determined. The sampled values of $\mathbf{F}(\mathbf{x})$ and the Taylor coefficients are related by the matrix equation

$$
\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
1 & x^{2} & x^{4} & x^{6} \\
1 & 4 x^{2} & 16 x^{4} & 64 x^{6} \\
1 & 9 x^{2} & 81 x^{4} & 729 x^{6}
\end{array}\right]\left[\begin{array}{c}
f_{0} \\
f_{2} \\
f_{4} \\
f_{6}
\end{array}\right]=\left[\begin{array}{c}
F(0) \\
F(x) \\
F(2 x) \\
F(3 x)
\end{array}\right] .
$$

Solving for the Taylor coefficients yields the equation.

$$
\left[\begin{array}{l}
f_{0} \\
f_{2} \\
f_{4} \\
f_{6}
\end{array}\right]=\frac{1}{360 x^{6}}\left[\begin{array}{cccc}
360 x^{6} & 0 & 0 & 0 \\
-490 x^{4} & 540 x^{4} & -54 x^{4} & 4 x^{4} \\
140 x^{2} & -195 x^{2} & 60 x^{2} & -5 x^{2} \\
-10 & 15 & -6 & 1
\end{array}\right]\left[\begin{array}{l}
F(0) \\
F(x) \\
F(2 x) \\
F(3 x)
\end{array}\right]
$$

The powers of x adjust for the sampling increment, x . To calculate the Taylor coefficients to fourth order only, use the following three point equation,

$$
\left[\begin{array}{l}
f_{0} \\
f_{2} \\
f_{4}
\end{array}\right]=\frac{1}{12 x^{4}}\left[\begin{array}{ccc}
12 x^{4} & 0 & 0 \\
-15 x^{2} & 16 x^{2} & -x^{2} \\
3 & -4 & i
\end{array}\right]\left[\begin{array}{l}
F(0) \\
F(x) \\
F(2 x)
\end{array}\right]
$$

These equations could be incoporated into a thin film design program to calculate a Taylor series representation of the coating performance suitable for use with the polarization aberration equations.

## The Fresnei Equations

The equations governing the reflection and refraction of light from interfaces take many forins depending on the nature of the interface. The discovery that light becomes polarized upon both refraction through and reflection from a transparent nonbirefringent material was made by E. Louis Malus (Malus, 1809). The relations governing this polarization phenomenon were first derived by A. Fresnel in 1821 in his theory of partial refraction (Fresnel, 1866a\&b). Fresnel postulated that light was
a transverse wave, and that polarized light was a manifestation of the two degrees of freedom associated with a transverse wave. Fresnel's hypothesis originally met with ridicule since the ether, which was assumed to be a fluid, would be incapable of sustaining transverse waves. None the less, the Fresnel equations provided a quantitative model for partial reflection. The Fresnel equations were rederived from electromagnetic theory first by P.Drude (Swindell, 1975, pg.8).

The Fresnel equations relate the relative amplitudes and phases of the components of an electromagnetic wave at planar interface. Figure 3 shows this configuration with the incident beam, the reflected beam and the transmitted beam. The wave vectors for the three beams are unit vectors aligned with the Poynting vectors of the three waves: $\hat{\mathbf{k}}_{0}$ for the incident beam, $\hat{\mathbf{k}}_{\mathrm{r}}$ for the reflected beam, and $\hat{\mathbf{k}}^{\prime}$ for the transmitted beam. The unit surface normal vector is $\hat{\mathrm{n}}$. The caret above a vector indicates a unit vector. All angles are measured from the surface normal. The incident piane wave has an angle of incidence, i. The angle of reflection equals the angle of incidence, while the wave vector lies in the plane of incidence on the opposite side of the surface normal from the incident wave vector. The reflected wavevector $\hat{\mathrm{k}}^{\prime}$ is determined from the vector law of reflection,

$$
\hat{k}^{\prime}=\hat{k}_{0}-2\left(\hat{k}_{0} \cdot \hat{n}\right) \hat{n} .
$$

The transmitted component propagates into the second medium at the angle of refraction, $i^{\prime}$, given by Snell's law,

$$
n \sin i=n^{\prime} \sin i^{\prime}
$$

One form of the vector law of refraction follows. Define a unit vector, $\hat{\mathrm{m}}$ in the plane of incidence and perpendicular to the normal

$$
\hat{m}=\frac{\hat{k}_{0}-\left(\hat{k}_{0} \cdot \hat{n}\right) \hat{n}}{\left|\hat{k}_{0}-\left(\hat{k}_{0} \cdot \hat{n}\right) \hat{n}\right|}
$$



Figure 3 Incident. Reflected and Refracted Waves at an Interface
At a dielectric interface with refractive indices n and $n$ ', the wavevectors for the incident, reflected and refracted beams are: $k_{0} k_{r}$ and $k$. The surface normal is $n$. The angles of incidence and refraction are $i$ and $i^{\circ}$.
where $|\vec{x}|$ is the norm of the vector $x$. From Snell's law:

$$
\begin{gathered}
\sin i^{\prime}=\frac{n}{n^{\prime}} \sqrt{1-\left|\hat{X}_{0} \cdot \hat{n}\right|^{2}}, \\
\cos i^{\prime}=\sqrt{1-\sin ^{2} i^{\prime}} .
\end{gathered}
$$

The refracted wave vector $\vec{k}^{\prime}$ is

$$
\hat{\mathbf{k}}^{\prime}=-\cos \mathrm{i}^{\prime} \hat{\mathbf{n}}+\sin \mathrm{i}^{\prime} \hat{m} .
$$

The refractive index of the two media are $n$ and $n^{\prime}$, where $n$ is associated with the incident medium. It is assumed that the media are isotropic and homogeneous; $n$ and $n^{\prime}$ are independent of orientation and location in the medium. Similarly, the interface is assumed to be isotropic and without any surface roughness. scratches, grating structure or any other features to break the symmetry.

The Fresnel amplitude transmission relations are well known and are not derived here. The derivation consists of matching the tangential $\overrightarrow{\mathbf{E}}$ and $\overrightarrow{\mathbf{H}}$ fields across the interface. Derivations are found in many references including Born and Wolf (1975, section 1.5.2), Marion (1965, section 6.3) and Stratton (1948, Chapt.9). The derivation in Stratton is noteworthy for the inclusion of the magnetic permittivity of the media, $\mu$, and an extended discussion of the results. The magnetic permittivity of optical glasses, reflective metals and coating materinls is usually so close to the magnetic permittivity of free space $\mu_{0}$ that it can be safely neglected.

The notation $a_{s}$ and $a_{p}$ refer to either the reflected or transmitted amplitude transmission coefficient, while $t_{s}$. $t_{p}$. $r_{s}$ and $r_{p}$ refer unambiguously to the transmitted and reflected components. The Fresnel amplitude equations are:

$$
t_{s}(i)=\frac{2 \cos i \sin i^{\prime}}{\sin \left(i+i^{\prime}\right)}=\frac{2 n \cos i}{n \cos i+n^{\prime} \cos i^{\prime}}
$$

$$
\begin{aligned}
& t_{p}(i)=\frac{2 \cos i \sin i^{\prime}}{\sin \left(i+i^{\prime}\right) \cos \left(i-i^{\prime}\right)}=\frac{2 n \cos i}{n^{\prime} \cos i+n \cos i^{\prime}} . \\
& r_{s}(i)=\frac{-\sin \left(i-i^{\prime}\right)}{\sin \left(i+i^{\prime}\right)}=\frac{n \cos i-n^{\prime} \cos i^{\prime}}{n \cos i+n^{\prime} \cos i^{\prime}} . \\
& r_{p}(i)=\frac{\tan \left(i-i^{\prime}\right)}{\tan \left(i+i^{\prime}\right)}=\frac{n^{\prime} \cos i-n \cos i^{\prime}}{n^{\prime} \cos i+n \cos i^{\prime}} .
\end{aligned}
$$

The Fresnel equations depend on the ratio of the indices, $n$ and $n^{\prime}$, but not on the values of the refractive indices individually. This relative refractive index ratio is defined as

$$
N=\frac{n}{\mathbf{n}^{\circ}}
$$

The Fresnel equations are equally valid for real $n$, corresponding to transparent media, or complex $n$, corresponding to absorbing media and metals.

The Fresnel equations are difficult to manipulate analytically because of their complicated form. The presence of the sum of two trigonometric functions in the denominator leads to very tecious integrals and derivatives. This complexity may have deterred cther attempts to include amplitude and polarization effects in an optical aberration theory. To put the Fresnel equations in a usable form for the polarization aberrations, the following Taylor series expansions have been calculated using the computer algebra program Macsyma (licensed by Symbolics Corp. Cambridge, MA).

The second order Taylor series expansions for the Fresnel amplitude coefficients about $\mathrm{i}=0$ are:

$$
\begin{aligned}
& t_{s}(i)=\frac{2 N}{N+1}+i^{2} \frac{N(N-1)}{N+1}, \\
& t_{p}(i)=\frac{2 N}{N+1}+i^{2} \frac{N^{2}(N-1)}{N+1}, \\
& r_{S}(i)=\frac{N-1}{N+1}-i^{2} \frac{N-1}{N(N+1)},
\end{aligned}
$$

$$
r_{p}(i)=\frac{N-1}{N+1}+i^{2} \frac{N-1}{N(N+1)}
$$

The following fourth order Taylor series terms have been obtained for the transmitted beam:

$$
\begin{aligned}
& t_{s, 4}(i)=i^{4} \frac{N\left(3 N^{3}+3 N^{2}-7 N+1\right)}{12(\hat{N}+i)} . \\
& t_{p, 4}(i)=i^{4} \frac{N^{2}\left(\frac{N}{}-1 / 2 N^{2}-6 N^{2}+5\right)}{12(N+1)} .
\end{aligned}
$$

## Intensity and Power Transmission Relations

The power in the reflected and transmitted beams for the $s$ and $p$ components of the light at an interface are determined as follows. Because the transmitted beam changes direction, the cross section of the transmitted beam in the plane of incidence (pplane) changes by a multiplicative factor

$$
K=\frac{\cos i^{\prime}}{\cos i}
$$

In the s plane, the transmitted beam cross section remains constant. Thus, a circular light beam retracts at a planar interface into an elliptical beam. The reflected beam at a planar interface does not change its cross section, so that for reflection

$$
K=1.0
$$

The power reflection and transmission coefficients are (Stratton 1948) :

$$
\begin{gathered}
R_{k}=\left|r_{k}\right|^{2} \\
T_{k}=\frac{n^{\prime} \cos i^{\prime}}{n \cos i}\left|t_{k}\right|^{2}
\end{gathered}
$$

where $k=s, p$. In the absence of absorption

$$
R_{k}+T_{k}=1.0
$$

The coefficients $R_{k}$ and $T_{k}$ are the fractional reflectance and transmittance of the energy in the beam, a ratio of incident to reflected or transmitted power. The power
transmission coefficient can be measured as a ratio using a radiometer where the test beams underfill the entrance pupil of the radiometer and all the energy is measured.

The intensity transmittances for the $s$ and $p$ components. $I_{s}$ and $I_{p}$ are defined as:

$$
I_{s}=A_{s}{ }^{2} \quad, \text { and, } I_{p}=A_{p}{ }^{2}
$$

The intensity transmittance is the ratio between the transmitted (or reflected) intensity and the incident intensity. It is what would be measured by a radiometer oriented normal to the beams whose entrance pupil is overfilled by the light, measuring the energy per area. The intensity transmittance, I, can exceed one when $\mathrm{K}>1$. For example, when refracting from glass into air at Brewster's angle, the transmitted beam has a smaller cross section than the incident beam. Since all of the p light is transmitted, the power transmission coefficient is one while intensity of the refracted p light is higher than the intensity of the incident light;

$$
\mathrm{T}_{\mathrm{p}}=1.0, \mathrm{I}_{\mathrm{p}}>1.0
$$

## Refraction, The Transmitted Beam

Consider light refracting at the interface between two dielectric, nonabsorbing media with refractive indices n and $\mathrm{n}^{\prime}$. When light is incident at this interface, it divides into two components, the transmitted and the reflected components. If neither medium has appreciable absorption then $n$ and $n^{\prime}$ are real. The second order amplitude transmission relations for the transmitted beam are:

$$
\begin{aligned}
& t_{s}(i)=\frac{2 N}{N+1}-i^{2} \frac{N(N-1)}{N+1} . \\
& t_{p}(i)=\frac{2 N}{N+1}-i^{2} \frac{N^{2}(N-1)}{N+1} .
\end{aligned}
$$

At normal incidence,

$$
t_{s}(\mathbf{i})=t_{p}(\mathbf{i}) .
$$

so there is no polarization on-axis. Since $n$ and $n^{\prime}$ are real, the amplitude transmission coefficients are real, not complex. Thus. there is no phase change on transmission for either the sor p component. Since both components are transmitted with zero phase change, there is no retardance induced at any angle of incidence.

There is a second order difference in the magnitude of the transmission
 refracting interface becomes a weak linear polarizer. Linearly polarized light is always refracted as linearly polarized light. Although the plane of polarization of linearly polarized light may rotate upon refraction due to the linear polarization, the ellipticity remains zero because of the absence of retardance.

The Taylor series expansion for the Jones matrix to second order in the angle of incidence and expressed in $S$ and $P$ coordinates is

$$
J(i)=\frac{2 n}{n+n^{\prime}}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]+i^{2} \frac{n\left(n-n^{\prime}\right)}{n^{\prime 2}\left(n+n^{\prime}\right)}\left[\begin{array}{cc}
n & 0 \\
0 & n^{\prime}
\end{array}\right] .
$$

The non-zero Taylor series coefficients to second order for the normalized $\mathbf{C}$ vector

$$
\vec{C}=\tau\left[1+d_{02} i^{2}, d_{12} i^{2}, 0,0\right]
$$

expressed in $S$ and $P$ coordinates are:

$$
\tau=d_{00}=\frac{2 N}{N+1}, d_{02}=\frac{(N-1)(N+1)}{4}, d_{12}=\frac{N(N-1)}{2} .
$$

The power transmittance, or fraction of the optical power transmitted across the refracting interface is given by the intensity transmittance relations:

$$
T_{s}=K A_{s}{ }^{2}, \text { and, } T_{p}=K A_{p}^{2}
$$

Figure 4 is a plot of the power transmission for an air to glass ( $\mathrm{N}=1.6$. upper pair of curves) and an air to germanium ( $\mathrm{N}=4.0$, lower pair) interface as a function of the angle of incidence. The $p$ amplitude transmission curves rise to a value of 1 at

Brewster's angle. The difference between the $s$ and $p$ curves is the polarization associated with the interface.

## Single Layer Dielectric Thin Films

Very few uncoated refracting interfaces are used in quality optical systems. Ccatings are commonly used to enhance the transmission through optical systems, as with antireflection and reflection enhancing coatings. Coatings are also applied to change the spectral distribution of light, to polarize the light and for many other purposes.

In this section, the polarization produced by single layer antireflection coatings will be examined. These coatings display the same polarization properties as multilayer coatings. The polarization properties of interest are the angle of incidence dependence of: the transmitted amplitude, phase, the linear polarization and the linear retardance. Isotropic thin films do not display circular polarization or circular retardance.

Figure 5 shows the model for the single layer thin film, an isotropic stratified planar layer of uniform thickness on a plane substrate. This problem is treated in Azzam and Bashara (1977, section 4.3). Born and Wolf (1975), Macleod (1974). The refractive indices of the incident medium, the thin film and the final medium are: $n_{0}, n_{1}$ and $n_{2}$. The two interfaces are assumed parallel. The angle of propagation in each medium is found from Snell's law,

$$
n_{0} \sin i_{0}=n_{1} \sin i_{1}=n_{2} \sin i_{2} .
$$

Also needed are.

$$
\cos i_{1}=\sqrt{1-\frac{n_{0}{ }^{2} \sin ^{2} i_{0}}{n_{1}{ }^{2}}}, \cos i_{2}=\sqrt{1-\frac{n_{0}{ }^{2} \sin ^{2} i_{0}}{n_{2}{ }^{2}}} .
$$

The effective thickness of the thin film varies with the angle of incidence. The film


Figure 4 Power Transmission of Dielectric Interfaces
The $s$ and $p$ power transmission as a function of angle of incidence is shown for two dielectric refracting interfaces, $N=\frac{\mathbf{n}^{\prime}}{\mathrm{n}}=1.6$ (upper pair), and $\mathrm{N}=4.0$.


Figure 5 A Single Layer Thin Film in Transmission
A single layer thin film of refractive index $\mathrm{n}_{1}$ and thickness d is located between two media with refractive indices $\mathrm{n}_{0}$ and $\mathrm{n}_{2} . \mathrm{i}_{0}, \mathrm{i}_{1}$ and $\mathrm{i}_{2}$ are the propagation angles.
phase thickness, $\beta$, in radians is

$$
\beta\left(\mathrm{n}_{0}, \mathrm{n}_{1}, \mathrm{n}_{2}, \mathrm{i}_{0}, \mathrm{~d}, \lambda\right)=\frac{2 \pi \mathrm{dn}_{1} \cos \mathrm{i}_{1}}{\lambda} .
$$

 interfaces are:

$$
\begin{aligned}
& r_{01 p}=\frac{n_{1} \cos i_{0}-n_{0} \cos i_{1}}{n_{1} \cos i_{0}+n_{0} \cos i_{1}}, \\
& r_{01 s}=\frac{n_{0} \cos i_{0}-n_{1} \cos i_{1}}{n_{0} \cos i_{0}+n_{1} \cos i_{1}}, \\
& r_{12 p}=\frac{n_{2} \cos i_{1}-n_{1} \cos i_{2}}{n_{2} \cos i_{1}+n_{1} \cos i_{2}}, \\
& r_{12 s}=\frac{n_{1} \cos i_{1}-n_{2} \cos i_{2}}{n_{1} \cos i_{1}+n_{2} \cos i_{2}}, \\
& t_{01 p}=\frac{2 n_{0} \cos i_{0}}{n_{1} \cos i_{0}+n_{0} \cos i_{1}} . \\
& t_{01 s}=\frac{2 n_{0} \cos i_{0}}{n_{0} \cos i_{0}+n_{1} \cos i_{1}}, \\
& t_{12 p}=\frac{2 n_{1} \cos i_{1}}{n_{2} \cos i_{1}+n_{1} \cos i_{2}}, \\
& t_{12 s}=\frac{2 n_{1} \cos i_{1}}{n_{1} \cos i_{1}+n_{2} \cos i_{2}} .
\end{aligned}
$$

The amplitude transmission coefficients for the single layer thin film system are:

$$
\begin{aligned}
& t_{p}\left(i, n_{0}, n_{1}, n_{2}, \beta\right)=\frac{t_{01 p} t_{12 p} e^{-j \beta}}{1+r_{01 p} r_{12 p} e^{-j 2 \beta}}, \\
& t_{s}\left(i, n_{0}, n_{1}, n_{2}, \beta\right)=\frac{t_{01 s} t_{12 s} e^{-j \beta}}{1+r_{01 s} r_{12 s} e^{-j 2 \beta}} .
\end{aligned}
$$

The reflection coefficients for the single layer thin film system are:

$$
\begin{aligned}
& r_{p}\left(i_{0}, n_{0}, n_{2}, n_{2}, \beta\right)=\frac{r_{01 p}+r_{12 p} e^{-j 2 \beta}}{1+r_{01 p} r_{12 p} e^{-j 2 \beta}}, \\
& r_{s}\left(i_{0}, n_{0}, n_{1}, n_{2}, \beta\right)=\frac{r_{01 s}+r_{12 s} e^{-j 2 \beta}}{1+r_{01 s} r_{12 s} e^{-j 2 \beta}} .
\end{aligned}
$$

The single layer thin film system displays both amplitude and phase variations with angle and wavelength. Sirce the solution is a set of separable equations for the $s$ and $p$ components, the eigenpolarizations are linearly polarized light aligned with the $s$ and $p$ planes. The eigenvalues are, in general, complex. Thus, there are amplitude and phase changes on reflection and refraction. These amplitude and phase changes are usually different for the $s$ and $p$ components, so that the coating displays polarization and retardance.

## Antireflection Coating Example

A particular antireflection coating, a magnesium fluoride ( $\mathrm{MgF}_{:}$) antireflection coating on glass, will be analyzed to gain insight into coating polarization behavior. The incident medium is air and the substrate is BK7 glass with a refractive index of 1.52. The refractive index of $\mathrm{MgF}_{2}$ will be assumed to be 1.38 , although large index variations can occur depending on the deposition conditions. Dispersion, the variation of refractive index with wavelength, is ignored.

At normal incidence, the optimum phase thickness of a $\mathrm{MgF}_{2}$ antireflection coating is $\beta=\pi / 2$, or a thickness, $d$, equivalent to a quarter of a wavelength of light in the film, $d=\lambda / 4 n_{1}$.

The following figures show the variation of polarization parameters (along the y axis) as a function of the angle of incidence (along the x axis) for the quarterwave


Figure 6 Power Transmission of a Quarter Wave Thin Film
The $s$ and $p$ power transmission of a quarter wave antireflection coating are plotted as a function of the angle of incidence.
$\mathrm{MgF}_{2}$ antireflection coating. Figure 6 shows the s and p power transmission coefficients as a function of the angle of incidence (in degrees.) Adding the coating has increased the transmission by reducing the normal incidence reflectance to 0.0125 as compared to 0.04 for the uncoated interface. As the angle of incidence increases, the p transmission increases to 1.0 at the Brewster angle for this coating. before decreasing. The s transmittance monotonically decreases. Figure 7 shows the phase change on transmission for the $s$ and $p$ components. The phase change is defined as the phase difference between the light incident at the first interface and the light at the substrate boundary. The absolute phase change at normal incidence is -90 degrees since the phase thickness of the medium is $\pi / 2$ radians. The differences between the $s$ phase change and the $p$ phase change, is very small for the quarter wave coating being not second order in the angle of incidence but fourth order.

A single layer coating can have a quarter wave of optical thickness at only a single wavelength. At other wavelengths, the polarization differences are larger. Figure 8 shows the $s$ and $p$ power transmission coefficients for three different wavelengths or, equivalently, three different phase thicknesses of $\mathrm{MgF}_{2}: \beta=\frac{\pi}{4}, \frac{\pi}{3}$ and $\frac{\pi}{2}$ (quarter wave). The further the coating thickness is from quarter wave, the lower the $s$ transmission is and the larger the polarization is. Figures 9 a and 9 b show the behavior of the phase for $\beta=\frac{\pi}{2}, \frac{2 \pi}{5}, \frac{\pi}{3}$ and $\frac{\pi}{4}$ coatings. Figure 9 a displays the mean phase change,

$$
\Delta(i)=\frac{\delta_{s}(i)+\delta_{p}(i)}{2} .
$$

The mean phase change is the wavefront aberration attributable to the coating. As i


Figure 7 The Phase Change of an Antireflection Coating
The $s$ and $p$ phase change upon transmission through a quarter wave antireflection coating are plotted as a function of the angle of incidence.


Figure 8 The $S$ and $P$ Transmission of an Antireflecting Coating
The S and P power transmission coefficient for transmission through a quarter wave (at $\lambda_{0}$ ) antireflection coating is plotted in degrees as a function of angle of incidence for $\lambda=\lambda_{0}, 1.25 \lambda_{0}, 1.5 \lambda_{0}$ and $2 \lambda_{0}$. The difference between $s$ and $p$ curves is the linear polarization as a function of angle of incidence, the quadratic portion of which is characterized by the real part of the $C$ vector coefficient $d_{12}$.

a) The mean phase change upon transmission through a quarter wave (at $\lambda_{0}$ ) antireflection coating is plotted in degrees as a function of angle of incidence for $\lambda=\lambda_{0}, 1.25 \lambda_{0}, 1.5 \lambda_{0}$ and $2 \lambda_{0}$. Curvature at the origin represents defocus in the wavefront introduced by the coating and is characterized by the imaginary part of the $C$ vector coefficient $d_{02}$.

b) The difference between $s$ and $p$ phase change, the linear retardance, upon transmission is plotted for the same coating and wavelengths. The quadratic portion of these curves is described by the imaginary part of the $C$ vector element $d_{12}$.

Figure 9 Phase Changes of an Antireflection Coating
approaches zero, the value of $\Delta(i)$ approaches zero as expected for the limit of no coating. Likewise the quadratic variation of $\Delta(i)$ is reduced as it approaches a constant value of zero for $\beta=0$. The quadratic variation of $\Delta(\mathrm{i})$ near $\mathrm{i}=0$ is defocus. For a spherical wave refracting at a spherical coated interface, this quadratic phase variation is the paraxial weak lensing due to the coating.

Figure $9 b$ shows the linear retardance, the difference in the $s$ and $p$ phase change upon transmission for these coatings

$$
\delta(i)=\delta_{s}(i)-\delta_{p}(i)
$$

This angle of incidence dependent retardance occurs even though none of the materials are birefringent. It arises because the transmitted light is the coherent superposition of beams which take multiple paths through the thin film. The small difference between the $s$ and $p$ amplitude coefficients means that different amounts of $s$ and $p$ light are present after $1,2,3$... bounces inside the coating. Since the transmitted light after different numbers of bounces has different relative phases, the sums of these beams have small retardances.

Without a coating , $\beta=0$. the retardance is zero for all angles of incidence. As $\beta$ increases, the retardance increases until it reaches a maximum near $\beta=\frac{\pi}{8}$ for small angles of incidence. Then the retardance decreases until $\beta=\frac{\pi}{4}$. For this quarter wave coating, there is no quadratic variation of retardance, only fourth and higher order dependence. It is a fortunate occurence that the quarter wave coating minimizes retardance for small angles of incidence. The most common antireflection coating in use has small retardance polarization aberration.

The objective is to calculate the polarization effects of coatings such as this quarter wave antireflection coating on light propagating through optical systems. To
do this with the polarization aberrations, the exact coating performance functions. such as are plotted in Figures 6. 7. 8, and 9, are replaced with approximate polynomial functions of the form (shown here for the complex amplitude transmission relations):

$$
\begin{aligned}
& a_{s}(i) \cong a_{s, 0}+a_{s,} 2^{i^{2}}+a_{s, 4} 4^{4^{4}}+\ldots=\tau\left(1+\left(d_{02}+d_{12}\right) i^{2}+\ldots\right) \\
& a_{p}(i) \cong a_{p, 0}+a_{p, 2} 2^{i^{2}}+a_{p, 4} 4^{i^{4}}+\ldots=T\left(1+\left(d_{02}-d_{12}\right) i^{2}+\ldots\right) .
\end{aligned}
$$

Polarization aberration thecry then relates these coefficients which characterize the coating ( $\tau, \mathrm{d}_{02}, \mathrm{~d}_{12}, \ldots$ ) to specific functions which characterize the variation of polarization and retardance associated with the optical system, the polarization aberration terms. Thus, the Taylor series coefficients form the bridge between thin film coatings and aberration theory.

## CHAPTER 4

## CASCADED WEAK POLARIZERS

## Introduction

In this chapter the properties of sequences of weak polarizers are developed and the following questions addressed. Can a series of linear polarizers be replaced with a single equivalent linear polarizer, a single polarizer who's Jones matrix equals the Jones matrix of the sequence? Likewise can a series of linear retarders be replaced with a single equivalent linear retarder?

Consider another question. Reflection, refraction and "fine" coatings behave as weak linear polarizers parallel to weak linear retarders. The eigenpolarizations of these interfaces are linear polarized light. Will the net polarization along a skew ray path through a sequence of such interfaces still be equivalent to a weak linear polarizer parallel to a weak linear retarder? That is, will the eigenpolarizations for the sequence be linearly polarized light?

The answer to all these questions is no. Cascaded polarizers are neither simple nor obvious. The $\mathbf{C}$ vectors are a method to simplify the problem and make tractable the understanding of series of weak polarizers.

The principle which summarizes this chapter is that "weak polarizers are very weakly order dependent." The behavior of a series of polarizers is order dependent. For weak polarizers, this order dependence is contained in the higher order terms.

## The Order Dependence of Pairs of Polarizers and Retarders

The order dependent polarization of several configurations of polarizers and retarders will be treated.

## Pairs of Linear Retarders

Consider two linear retarders normal to collimated light propagating along the $z$ axis. Let the fast axis of retarder $R$ be parallel with the $x$ axis. Retarder $R^{\prime}$ has its fast axis arbitrarily oriented in the $x-y$ plane. The normalized $C$ vectors for $R$ and $R^{\prime}$ are:

$$
\overrightarrow{\mathbf{R}}=\tau\left[1, j b_{1}, 0,0\right] \quad \text {, and, } \quad \vec{R}^{\prime}=\tau^{\prime}\left[1, j b_{1}^{\prime}, j b_{2}^{\prime}, 0\right]
$$

The $C$ vectors have been normalized by factoring out $c_{0}$ as $\tau$. Let $\overrightarrow{\mathrm{D}}$ be the C vector for the combination of retarders in which the light passes first through $R$ then $R^{\prime}$. Similarly, $\vec{E}$ is the $C$ vector for the reverse combination, $R^{\prime}$ then $R$. The $C$ vectors for these combinations are calculated using the "basis coupling matrix" $K$ yielding:

$$
\begin{aligned}
& =\tau^{\prime} \tau\left[1-b_{1} b_{1}{ }^{\prime}, j\left(b_{1}+b_{1}{ }^{\prime}\right), j b_{2}^{\prime}, j b_{1} b_{2}{ }^{\prime}\right] \text {. }
\end{aligned}
$$

and,

$$
\begin{aligned}
\vec{E}=\mathbf{K}(\overrightarrow{\mathrm{R}}) \overrightarrow{\mathrm{R}}^{\prime} & =\pi \tau^{\prime}\left[\begin{array}{cccc}
1 & j b_{1} & 0 & 0 \\
\mathrm{jb} & 1 & 0 & 0 \\
0 & 0 & 1 & \mathrm{~b}_{1} \\
0 & 0 & -\mathrm{b}_{1} & 1
\end{array}\right]\left[\begin{array}{c}
1 \\
j b_{1}^{\prime} \\
j b_{2}^{\prime} \\
0
\end{array}\right] \\
& =\pi \tau^{\prime}\left[1-\mathrm{b}_{1} \mathrm{~b}_{1}^{\prime}, j\left(\mathrm{~b}_{1}+\mathrm{b}_{1}^{\prime}\right), j b_{2}^{\prime},-j b_{1} b_{2}^{\prime}\right]
\end{aligned}
$$

Since the $c_{3}$ elements of $\vec{D}$ and $\vec{E}, D_{3}$ and $E_{3}$, are nonzero. the combinations $\vec{D}$ and $\vec{E}$ are not pure linear retarders. Since the element $c_{3}$ is imaginary, it represents the
presence of a circular retardance in these combinations. All the $\mathbf{C}$ vector elements of $\vec{D}$ and $\vec{E}$ are equal except for $D_{3}$ and $E_{3}$ for which

$$
D_{3}=-E_{3} .
$$

Thus, the order dependence only affects the sign of the circular retardance introduced; the linear retardance is order independent.

For a sequence of linear retarders oriented at either zero or 90 degrees with respect to the first retarder, there is no order dependence; the Jones matrices all commute.
$R$ and $R^{\prime}$ are weak linear retarders if

$$
b_{1}, b_{1}{ }^{\prime}, b_{2}^{\prime} \ll 1 .
$$

Consider this weak linear retardance as a first order perturbation to an identity matrix. The circular retardance term, $c_{3}$, and the change to the zero'th component, $\Delta c_{0}$. are second order perturbations since:

$$
c_{3}=b_{1} b_{2}^{\prime} \ll b_{1}, b_{2}^{\prime} \text {, and, } \Delta c_{0}=b_{1} b_{1}^{\prime} \ll b_{1}, b_{1}^{\prime} .
$$

For pairs of weak linear retarders, the circular retardance introduced is very weak and is the oniy order dependent term present. When second order terms are neglected, the resulting first order $C$ vectors equal the vector sum of $R$ and $R^{\prime}$. Let the order operator $\mathrm{O}\{f(\mathrm{x}), \mathrm{n}\}$ return the terms in $\mathrm{f}(\mathrm{x})$ of order less than or equal to n . Then.

$$
\mathrm{O}\{\overrightarrow{\mathrm{D}}, 1\}=\pi \tau^{\prime}\left[1, \mathrm{j}\left(\mathrm{~b}_{1}+\mathrm{b}_{2}{ }^{\prime}\right), \mathrm{j} \mathrm{~b}_{2}^{\prime}, 0\right]=\overrightarrow{\mathrm{R}}+\overrightarrow{\mathrm{R}}^{\prime} .
$$

and,

$$
\mathrm{O}\{\overrightarrow{\mathrm{E}} .1\}=\pi \tau^{\prime}\left[1, \mathrm{j}\left(\mathrm{~b}_{1}+\mathrm{b}_{1}^{\prime}\right), \mathrm{jb}_{2}^{\prime}, 0\right]=\overrightarrow{\mathrm{R}}+\overrightarrow{\mathrm{R}}^{\prime} .
$$

## Example

Consider two half wave plates with their fast axes oriented at $45^{\circ}$ to each other. The C vectors are:

$$
\vec{R}=[0,1,0,0], \text { and, } \vec{R}^{\prime}=[0,0,1,0]
$$

The C vectors for the combinations are:

$$
\begin{aligned}
\vec{D} & =\mathbf{K}(\vec{R}) \vec{R}=[0,0,0,-j] \\
\vec{E} & =\mathbf{K}(\vec{R}) \vec{R}^{\prime}=[0,0,0, j] .
\end{aligned}
$$

These are the $\mathbf{C}$ vectors for half wave circular retarders. This is the maximum possible (total) polarization coupling. It is a convenient way to assemble a circular retarder.

## Example

Consider a weak linear retarder oriented along the $\mathbf{x}$ axis

$$
\overrightarrow{\mathbf{R}}=[1, j \delta, 0,0]
$$

and ancther at $45^{\circ}$.

$$
\vec{R}^{\prime}=[1, j \epsilon, j \epsilon, 0] .
$$

Then:

$$
\begin{aligned}
& \vec{D}=\mathbf{K}\left(\overrightarrow{R^{\prime}}\right) \vec{R}=[1-\delta \epsilon, j(\delta+\epsilon), j \epsilon, j \delta \epsilon] . \\
& \vec{E}=\mathbf{K}(\vec{R}) \vec{R}^{\prime}=[1-\delta \epsilon, j(\delta+\epsilon), j \epsilon,-j \delta \epsilon] .
\end{aligned}
$$

The corresponding Jones matrix is

$$
J=\left[\begin{array}{cc}
1-\delta \epsilon+\mathrm{j}(\delta+\epsilon) & \mathrm{j} \epsilon+\delta \epsilon \\
\mathrm{j} \epsilon_{7} \delta \epsilon & 1-\delta \epsilon-\mathrm{j}(\delta+\epsilon)
\end{array}\right]
$$

The upper sign of the $\neq$ and $\pm$ terms refers to $\overrightarrow{\mathrm{D}}$ and the lower to $\overrightarrow{\mathrm{E}}$. The $\mathrm{j} \delta \epsilon \sigma_{3}$ term is the induced circular retardance.

## Pairs of Linear Polarizers

Consider two weak linear polarizers, $\overrightarrow{\mathrm{L}}$ and $\overrightarrow{\mathrm{L}}^{\prime}$, in series. Their normalized C vectors are:

$$
\overrightarrow{\mathrm{L}}=\tau\left[1, \mathrm{a}_{1}, 0,0\right], \quad \text { and, } \quad \overrightarrow{\mathrm{L}}^{\prime}=\tau^{\prime}\left[1, a_{1}^{\prime}, a_{2}^{\prime}, 0\right]
$$

The C vectors for the two sequences are:

$$
\overrightarrow{\mathrm{D}}=\mathbf{K}\left(\overrightarrow{\mathrm{L}}^{\prime}\right) \overrightarrow{\mathrm{L}}=\tau^{\prime} \tau\left[1-\mathrm{a}_{1} \mathrm{a}_{1}^{\prime}, \mathrm{a}_{1}+a_{1}^{\prime}, a_{2}^{\prime},-j a_{1} a_{2}^{\prime}\right] \text {. }
$$


#### Abstract

$\vec{E}=\mathbf{K}(\vec{L}) \vec{L}^{\prime}=T \tau^{\prime}\left[1-a_{1} a_{1}^{\prime}, a_{1}+a_{1}{ }^{\prime}, a_{2}^{\prime}, j a_{1} a_{2}^{\prime}\right] \quad$. The presence of a nonzero $c_{3}$ component: ine $\pm j a_{1} a_{2}$ ' term. signifies that these combinations of linear polarizers do not act as pure linear polarizers. This $\mathrm{c}_{3}$ term describes a circular retardance component. Since circular retarders rotate linear polarized light. it is natural for circular retarders to appear in expressions involving sequences of linear polarizers when these sequences produce an overall rotation of polarization upon the incident light. In general, the eigenpolarizations of combinations of linear polarizers will be elliptically polarized light. The circular retardance associated with sequences of weak linear polarizers is a second order term which reaches a maximum when the linear polarizers are criented at $45^{\circ}$ with respect to each other.


## Pairs of Linear Polarizers and Linear Retarders

The property that pairs of linear polarizers or pairs of linear retarders couple into circular retardance is a consequence of the Pauli spin matrix identity

$$
\sigma_{1} \sigma_{2}=-\sigma_{2} \sigma_{1}=\mathrm{j} \sigma_{3} .
$$

This relation governs the coupling between linear "polarizers", either polarizers or retarders, in series at arbitrary angles with respect to each other. If either two linear polarizers are in series or two linear retarders are in series at angles other than $\theta=0^{\circ}, 90^{\circ}$ with respect to each other, a circular retardance term is produced.

The closely related cases of
(1) a linear polarizer followed by a linear retarder at $\theta \neq 0^{\circ}, 90^{\circ}$, or,
(2) a linear retarder followed by a linear polarizer at $\theta \neq 0^{\circ}, 90^{\circ}$.
both couple into circular polarization (not retardance) through the same Pauli spin matrix identities. These cases involve the terms:

$$
\text { (1) } \quad\left(\sigma_{0}+j b \sigma_{2}\right)\left(\sigma_{0}+a \sigma_{1}\right)=\left(\sigma_{0}+a \sigma_{1}+j b \sigma_{2}+a b \sigma_{3}\right) \text {. }
$$

$$
\text { (2) }\left(\sigma_{0}+a \sigma_{1}\right)\left(\sigma_{0}+j b \sigma_{2}\right)=\left(\sigma_{0}+a \sigma_{1}+j b \sigma_{2}-a b \sigma_{3}\right) \text {, }
$$

where the real $\sigma_{3}$ term is the induced circular polarization.

## Pairs Involving Circular Polarizers and Circular Retarders

Two other Pauli spin matrix identities involving circular polarizers and circular retarders also derive from the Fauli algebra:

$$
\sigma_{2} \sigma_{3}=-\sigma_{3} \sigma_{2}=\mathrm{j} \sigma_{1} \quad \text {, and. } \quad \sigma_{3} \sigma_{1}=-\sigma_{1} \sigma_{3}=\mathrm{j} \sigma_{2} .
$$

These relations govern the coupling between linear and circular polarization elements. A linear polarization element oriented at $\alpha$ and a circular polarization element couple into a linear component at $\alpha+45^{\circ}$.

Pairs involving only circular polarizers and/or circular retarders do not couple into linear polarization or linear retardance since

$$
\sigma_{3} \sigma_{3}=\sigma_{0} .
$$

## Summary of Polarization Couplings

All the polarization couplings are summarized in the basis coupling matrix. $\mathbf{K}(\vec{C})$, introduced in Chapter Two. This matrix, used there for calculating cascaded polarizers in the $\mathbf{C}$ vector notation, is,

## Instrumental Polarization For Ray Paths Through Optical Systems

In this section the Jones matrix representing the instrumental polarization for light propagating along a ray path through an optical system is derived. Results are also given for the instrumental polarization associated with paraxial rays as functions
of the Taylor series of the $C$ vectors representing the optical interfaces. The notation used in this section is compiled in Table 10.

## General Case

Consider an optical system with $\mathbf{Q}$ optical interfaces numbered in the order encountered from $q=1$ to $Q$. No symmetry regarding the optical configuration is assumed. Light propagates along a specified ray path such as would be calculated by an optical ray trace calculation. At each interface some polarization effects are introduced due to differences in the optical constants across the interface. In addition, polarization and retardance are associated with the ray path between interfaces due to optically active crystals, dichroism, birefringence, gradient index materials or other polarizing mechanisms.

For this ray, the Jones matrix representing the ray intercept at surface $q$ is denoted by $\mathrm{J}_{\mathbf{q}}$. The Jones matrix associated with the optical path between surfaces q and $\mathrm{q}+1$ is denoted by $\mathrm{L}_{\mathrm{q}}$. The instrumental polarization, J , associated with this ray path is

$$
\mathbf{J}=\prod_{\mathrm{u}=\mathbf{Q},-1}^{1} \mathbf{L}_{\mathbf{q}} \mathbf{J}_{\mathbf{q}}
$$

This is the most general case for the instrumental polarization associated with an optical ray path.

Light refracting into anisotropic media usually divides into two beams (an ordinary and extraordinary ray) in the second medium. The last equation is then applied separately to each beam.

## Homogeneous Optical Systems

A homogeneous interface has optical properties independent of spatial coorcinates on the interface. The Jones matrix and $C$ vectors are functions onl; of

## TABLE 10

## Notation for Chapter 4

| C | C vector |
| :--- | :--- |
| $c_{k}$ | $d_{k}$ coefficients rotated into arbitrary plane of incidence |
| $d_{k}$ | Normalized $C$ vector components in s-p coordinates |
| i | Angle of incidence |
| j | $\sqrt{ }-1$ |
| J | Jones matrix |
| $\mathbf{k}$ | Pauli spin matrix index: $0,1,2,3$ |
| $\mathbf{K}$ | Basis coupling matrix |
| $\boldsymbol{l}$ | Length of a ray segment |
| $\mathbf{L}$ | Jones matrix associated with a ray segment |
| $\mathbf{q}$ | Surface index |
| $\mathbf{Q}$ | Total number of surfaces |
| $\alpha, \beta$ | Direction cosines |
| $\boldsymbol{\theta}$ | Orientation of the plane of incidence |
| $\rho_{\mathrm{k}}$ | Absorption or polarization coefficient |
| $\boldsymbol{\sigma}$ | Pauli spin matrix |
| $\tau$ | Normal transmittance |
| $\phi_{\mathrm{k}}$ | Phase or retardance coefficient |

Subscript Ordering: k, 1, q .

For example, $d_{123}$, is the coefficient for:
the $\sigma_{1}$ polarization basis state.
that is second order in the angle of incidence Taylor series, $\mathbf{i}^{2}$. for the third interface.
the angle of incidence, plane of incidence, and optical properties of the interface media,
$J\left(i, \theta, q, q^{\prime}\right)$ and $\left.{ }^{( } \mathbf{( i}, \theta, q, q^{\prime}\right)$.
A optical surface or thin film coating with spatial variations or localized defects is not homogeneous. A fine diffraction grating is not homogeneous at the microscopic scale, but might be considered as homogeneous at a larger scale.

Likewise, a homogeneous medium has optical properties independent of spatial coordinates. An anisotropic crystalline medium is hornogeneous if it consists of a single crystal. The refractive index varies with direction but not with position. The instrumental polarization associated with a ray path in a homogeneous medium depends only on the optical constants and the direction and path length of the ray. If $\alpha$ and $\beta$ are two of the direction cosines, then the Jones matrix and $C$ vector associated with a ray path of length $l$ through a homogeneous medium are

$$
\mathrm{L}(\alpha, \beta, \ell) \quad \text { and, } \quad \overrightarrow{( }(\alpha, \beta, \ell) .
$$

A homogeneous optical system is composed entirely of homogeneous interfaces between homogeneous media. The Jones matrix associated with a ray path through a homogeneous system is,

$$
\mathbf{J}=\prod_{\mathbf{q}=\mathbf{Q},-1}^{1} \mathbf{L}_{\mathbf{q}}(\alpha, \beta, \ell) \mathbf{J}_{\mathbf{q}}(\mathrm{i}, \theta) .
$$

## Radially Symmetric Systems of Lenses, Mirrors and Coatings

This section develops the the polarization properties of optical systems comprised of lenses, mirrors and "fine" coatings for ray paths near the optical axis (the paraxial regime.) A radially symmetric optical system has an axis of symmetry. the optical axis. It is assumed that the optical elements and materials used in transmission are highly transparent and nonpolarizing, as is usual in lenses. The
polarization contribution, the L's, from the path lengths through highly transparent elements is often assumed small relative to the polarization arising at the interfaces and can be neglected when

$$
\mathbf{L}(\alpha, \beta, \ell) \cong \tau\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] .
$$

For this paraxial development to be accurate, it is necessary that the angles of incidence are small enough that the polarization associated with the interfaces is adequately approximated by a second order expansion of the $\mathbf{C}$ vector as a function of the angle of incidence. For an uncoated lens or mirror, this approximation is generally valid for $\mathrm{i}<30^{\circ}$. Caicuiation of the fourth and higher order coefficients allows estimation of the accuracy of these second order equations. The paraxial region for this polarization analysis is typically orders of magnitude larger than the paraxial region of geometrical optics (where the fourth and higher order wavefront aberrations are negligible.)

Homogeneous and isotropic interfaces do not display polarization at normal incidence. There is only an amplitude and phase change which is represented by the complex number, $\tau$, the normal amplitude transmittance. An isotropic interface such as a lens, mirror or coating has a $\mathbf{C}$ vector Taylor series in s-p ccordinates $(\theta=0)$ of the form

$$
\overrightarrow{\mathrm{d}}(\mathrm{i}, \theta=0)=\tau[1,0,0,0]+\mathrm{i}^{2} \tau\left[\mathrm{~d}_{02}, \mathrm{~d}_{12}, 0,0\right]+\mathrm{i}^{4} \tau\left[\mathrm{~d}_{04}, \mathrm{~d}_{14}, 0,0\right]+\ldots .
$$

For an arbitrary orientation $\theta$ of the plane of incidence, the $C$ vector is

$$
\mathrm{O}_{\mathrm{i}, \theta},=\tau[1,0,0,0]+\mathrm{i}^{2} \tau\left[\mathrm{c}_{02}(\theta), \mathrm{c}_{12}(\theta), \mathrm{c}_{22}(\theta), 0\right]+\mathrm{i}^{4} \tau\left[\mathrm{c}_{04}(\theta), \mathrm{c}_{14}(\theta), \mathrm{c}_{24}(\theta), 0\right]+\ldots .
$$

where the c's are determined from the d's by a rotational change of basis. Since homogeneous and isotropic interfaces do not display circular retardance or circular
polarization. $\sigma_{3}$ is not included to simplify the mathematics. The inclusion of circular polarization and circular retardance effects is straightforward.

## The C Vector for a Paraxial Ray

Consider a paraxial ray path through an optical system from surfaces $\mathrm{q}=1$ to $Q$ with angles of incidence, $\mathbf{i}_{\mathbf{q}}$, and orientations of the plane of incidence, $\theta_{\mathbf{q}}$. The Jones vector associated with the axial ray (down the optical axis.) $\mathrm{i}_{\mathrm{q}}=0$ for all q , is

$$
J=\tau_{Q} \sigma_{0} \tau_{Q-1} \sigma_{0} \ldots \tau_{2} \sigma_{0} \tau_{1} \sigma_{0}=\Upsilon \sigma_{0} .
$$

where

$$
\Upsilon \equiv \prod_{\mathbf{q}=1}^{\mathbf{Q}} \tau_{\mathbf{q}}
$$

The complex amplitude transmittance down the axis, $\Upsilon$, is the product of the normal incidence complex amplitude transmittances at each surface.

The Jones matrix associated with a ray at interface $q$ can be expressed in terms of the expansion of the interface Jones matrix as

The Jones matrix associated with the entire paraxial ray path resulting from keeping terms to second order at each interface is ( x indicates multiplication carried onto the next line)

$$
\begin{gathered}
J=\tau_{\mathbf{q}}\left(\sigma_{0}+i^{2}\left(c_{0,2, Q^{0}}^{\left.\left.\sigma_{0}+c_{1,2, Q} \sigma_{1}+c_{2,2, Q^{2}}^{\sigma_{2}}\right)\right)} x\right.\right. \\
\tau_{Q-1}\left(\sigma_{0}+i_{Q-1}^{2}\left(c_{0,2, Q-1} \sigma_{0}+c_{1,2, Q-1} \sigma_{1}+c_{2,2, Q-1} \sigma_{2}\right)\right) \quad x \\
\ldots \quad \tau_{2}\left(\sigma_{0}+i_{2}^{2}\left(c_{0,2,2} \sigma_{0}+c_{1,2,2} \sigma_{1}+c_{2,2,2} \sigma_{2}\right)\right) \quad x \\
\tau_{1}\left(\sigma_{0}+i_{1}^{2}\left(c_{0,2,1} \sigma_{0}+c_{\left.\left.1,2,1 \sigma_{1}+c_{2,2,1} \sigma_{2}\right)\right)}\right.\right.
\end{gathered}
$$

Associated with each interface are four terms. Performing the multiplications leads to $4^{Q}$ terms, all in even powers in $i$. Collecting terms of equal power in $i$, there is one term at zero'th order and $3 Q$ terms at second order. If $i$ is assumed small, the large number of higher order terms are of diminishing importance. The expression for $\mathbf{J}$ through second order is

$$
\begin{aligned}
J^{(0)}+J^{(2)}=\Upsilon \sigma_{0} & +\Upsilon \sigma_{0} \sum_{q=1}^{Q} i_{q}^{2} c_{0,2, q}+\Upsilon \sigma_{1} \sum_{q=1}^{Q} i_{q}^{2} c_{1,2, q} \\
& +\Upsilon \sigma_{2} \sum_{q=1}^{Q} i_{q}^{2} c_{2,2, q} .
\end{aligned}
$$

Since no polarization or retardance was assumed on axis, the contributions to the second order polarization for this ray are just sums of contributions from each surface. The multiplication taking place at second order is of the form (dropping the $c$ 's and using the $\sigma_{1}$ terms as an example)

$$
\mathrm{i}^{2} \mathrm{Q}_{1} \sigma_{1} \sigma_{0} \sigma_{0} \sigma_{0} \ldots \sigma_{0}+\sigma_{0} \mathrm{i}^{2} Q_{-1} \sigma_{1} \sigma_{0} \sigma_{0} \sigma_{0} \ldots \sigma_{0}+\ldots+\sigma_{0} \sigma_{0} \sigma_{0} \ldots \mathrm{i}_{2}^{2} \sigma_{1} \sigma_{0}+\sigma_{0} \sigma_{0} \sigma_{0} \ldots \sigma_{0} \mathrm{i}_{1}^{2} \sigma_{1}
$$

When the elements display no polarization or retardance at normal incidence, there is no order dependence in the second order terms. Only one non-identity matrix term occurs in each second order matrix product. The second order polarization associated with the paraxial ray path is obtained by a simple summation of second order polarization contributions at each intercept.

The lowest order polarization where interface order dependence occurs in homogeneous and isotropic systems is at fourth order. The fourth order expression involves all combinations of two second order polarization terms as well as single fourth order polarization terms occurring in isolation. The fourth order Jones matrix terms are
$J^{(4)}=T \sigma_{0} \sum_{q=1}^{Q} i_{q}^{i} c_{0.4, q}$
$+\Upsilon \sigma_{0} \sum_{q=1}^{Q} \sum_{q^{\prime}=1}^{q-1} i_{q^{2}} i^{2} q^{\prime}\left(c_{0,2, q^{\prime}} c_{0,2, q^{\prime}}+c_{1,2, q^{c}, 2, q^{\prime}}\right.$
$+\mathrm{c}_{2,2, \mathrm{q}^{\mathrm{c}} 2,2, \mathrm{q}^{\prime}}$ )
$+r \sigma_{1} \sum_{\mathrm{q}=1}^{\mathrm{Q}} \mathrm{i}_{\mathrm{q}}^{\mathrm{f}} \mathrm{c}_{1.4 . \mathrm{q}}$
$+\Upsilon \sigma_{1} \sum_{q=1}^{Q} \sum_{q^{\prime}=1}^{q-1} i_{q^{2}} i^{2},\left(c_{0,2, q^{c}}{ }^{c}, 2, q^{\prime}+c_{1,2, q^{c}} c_{0,2, q^{\prime}}\right)$
$+\mathrm{r}_{\mathrm{o}} \sum_{\mathrm{q}=1}^{\mathrm{Q}} \mathrm{i}_{\mathrm{q}}^{\mathrm{t}} \mathrm{c}_{2,4, \mathrm{q}}$
$+\Upsilon \sigma_{2} \sum_{q=1}^{Q} \sum_{q^{\prime}=1}^{q-1} i_{q^{2} i^{\prime}}{ }^{\prime}\left(c_{0,2, q^{c}}{ }^{c} 2,2, q^{\prime}+c_{2,2, q^{\prime}} c_{0,2, q^{\prime}}\right)$
$+\Upsilon \sigma_{3} \sum_{q=1}^{Q} \sum_{q^{\prime}=1}^{q-1}{ }^{i} \mathbf{i}^{2} i^{2}{ }^{\prime}\left(c_{1,2, q^{c}}{ }^{c}, 2, q^{\prime}-c_{2,2, q^{\prime}} c_{1,2, q^{\prime}}\right)$.
The circular polarization and circular retardance terms $\left(\sigma_{3}\right)$ are the only order dependent elements at fourth order. These terms, arising from the multiplication of $\sigma_{1}$ and $\sigma_{2}$, always occur in the form of the quantum mechanical commutator.

## CHAPTER 5

## RESULTS FROM GEOMETRICAL OPTICS

## Introduction

This chapter contains the material from geometrical optics required for the development of the polarization aberrations. It presents several results regarding the angles of incidence of paraxial skew rays which are derived in Appendix D. Then, the wavefront aberrations are defined and formulated in terms of Jones matrices. In this formulation the wavefront aberrations are recognized as one of eight subsets of the polarization aberralions. The other seven subsets are introduced in Chapter 6.

## Paraxial Optics Summary

The polarization aberrations are a description of the polarization behavior of an optical system expressed as an expansion in ray coordinates about the centers of the object and pupil. It is appropriate and convenient to obtain the derivations from a paraxial ray trace; appropriate, because understanding the instrumental polarization near the center of the pupil and image is key to understanding instrumental polarization in general; convenient because the paraxial ray trace is linear and easy to manipulate.

The paraxial ray trace is discussed in many optics books. The results required here are expressions for the angles and plane of incidence of paraxial rays. Chapter Three of Lens Design Fundamentals by R. Kingslake provides a explanation of paraxial ray trace methods. Other treatments include: Gerrard and Burch (1975, Chapter 2). Stavroudis (1982. Section 1.2), and Welford (1974, Chapter 2.)

Appendix D. "Paraxial Skew Rays", contains an extended treatment of the material in this review, particularly derivations of the expressions for the angle of incidence and its orientation. Also included are relations for calculating the paraxial angles of incidence from a paraxial ray trace.

## Coordinate System

The coordinate system used here is a normalized right handed coordinate system. The $z$ axis is the optical axis of a rotationally symmetric optical system. Light initially travels in the direction of increasing $\mathbf{z}$. Figure 10 shows the notation.

For a rotationally symmetric system, the object can be located on the $y$ axis without loss of generality. The object coordinate H is normalized such that

$$
H=0
$$

in the center of the field (on the optical axis) and

$$
\mathrm{H}=1
$$

at the nominal edge of the field of view.
The location where a ray strikes the entrance pupil is specified by the polar pupil coordinates $\rho$ and $\phi$. $\rho$ is normalized such that at the edge of a circular pupil

$$
\rho=1 .
$$

$\phi$ is defined here as it is in much of geometric optics, and in defiance to most analytical geometry, as being zero on the $y$ axis and increasing counterclockwise. Normalized Cartesian pupil coordinates x and y are occasionally used. They are defined as:

$$
x=-\rho \sin \phi, \text { and, } y=\rho \cos \phi .
$$

## Subscripts

Subscripts c and m refer to quantities associated with the chief and marginal rays in the $y-z$ plane. Subscript $q$ refers to the quantity at the $q^{\text {th }}$ interface in the


Figure 10 Coordinate System for Paraxial Rays
Rays through an optical system are characterized by ray coordinates at the object and entrance pupil. H is the normalized object coordinate; $\rho$ is the normalized pupil radius, $\phi$ is the polar angle in pupil measured counterclockwise from y axis. Alternatively x and y are normalized Cartesian pupil coordinates. The chief and marginal rays are also shown.
system. Where a set of expressions refers to a single surface, the surface subscript will be omitted. Subscript $e$ is the quantity evaluated at the system stop (mnemonic - entrance pupil.)

## The Paraxial Angle of Incidence

Expressions for the angle of incidence $i$ and the orientation of the plane of incidence $\theta$ of a ray at a given surface $q$ are expressed in terms of the marginal ( $\mathrm{i}_{\mathrm{m}, \mathrm{q}}$ ) and chief ray ( $\mathrm{i}_{\mathrm{c}, \mathrm{q}}$ ) angles of incidence at that surface. Details of the derivation are in Appendix D.

Assume that a paraxial ray trace has been performed for an optical system and that $i_{m, q}$ and $i_{c, q}$ have been calculated. A ray from normalized object coordinate $H$ which passes through pupil coordinates $\rho$ and $\phi$ has an angle of inciuance $i_{q}$ and orientation of the plane of incidence $\theta_{q}$ at surface $q$ equal to:

$$
\begin{gathered}
\mathrm{i}_{\mathrm{q}}=\sqrt{\mathrm{H}^{2} \mathrm{i}_{\mathrm{c}, \mathrm{q}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}, \mathrm{q}_{\mathrm{m}, \mathrm{q}}}+\rho^{2} \mathrm{i}_{\mathrm{m}, \mathrm{q}}^{2}} \\
\theta_{\mathrm{q}}=\arcsin \left[\frac{\rho \sin \phi \mathrm{i}_{\mathrm{m}}}{\left|\mathrm{i}_{\mathrm{q}}\right|}\right] .
\end{gathered}
$$

Figure 11 shows the paraxial angle and plane of incidence for three field angles. The position of the center of a line represents position in the circular aperture. The magnitude of the angle of incidence is represented by the length of the lines. The orientation of the lines corresponds to the orientation of the plane of incidence. Off axis the pattern is a simple translated version of the on axis pattern.

## The Wavefront Aberrations

In this section the monochromatic wavefront aberrations of a radially symmetric optical system are stated. The wavefront aberrations are a set of basis functions which describe the shape of an optical wavefront, usually in the exit pupil


Field Angle. $2 \mathrm{H}_{\mathrm{O}}$


## Field Angle, $\mathrm{HO}_{0}$



Field Anale. In Axis

Figure 11 Paraxial Angles of Incidence
The angle and plane of incidence for paraxial rays at a spherical surface are represented by the length and orientation of lines for an on-axis and two off-axis objects.
of an optical system. The wavefront aberrations are usually defined with respect to a reference sphere centered on the Gaussian image point (the paraxial image). For radially symmetric optical systems, the transmitted wavefront from object points ( $\mathrm{G}, \mathrm{H}$ ) at an equal radius from the center of the object.

$$
|\overrightarrow{\mathrm{H}}|=\mathrm{G}^{2}+\mathrm{H}^{2}=\text { constant },
$$

have identical form and differ only in orientation. The object can be located on the y axis without loss of generality. Following Shack (1982), the wavefront is a runction of only $\vec{H} \cdot \vec{H}, \vec{H} \cdot \vec{\rho}$ and $\vec{\rho} \cdot \vec{\rho}$ and can be expressed as

$$
\begin{aligned}
W(H, \rho, \phi) & =W(\vec{H} \cdot \vec{H}, \vec{\beta} \cdot \vec{\rho} \cdot \overrightarrow{\mathrm{H}} \cdot \vec{\rho}) \\
& =\sum_{\mathrm{Q}=0}^{\infty} \sum_{\mathrm{R}=0}^{\infty} \sum_{\mathrm{S}=0}^{\infty} W_{\mathrm{Q}, \mathrm{R}, \mathrm{~S}}\left(\overrightarrow{(\vec{H} \cdot \overrightarrow{\mathrm{H}})^{\mathrm{Q}}(\vec{\beta} \cdot \vec{\rho})^{R}(\overrightarrow{\mathrm{H}} \cdot \vec{\rho})^{S}}\right. \\
& =\sum_{\mathrm{Q}=0}^{\infty} \sum_{\mathrm{R}=0}^{\infty} \sum_{\mathrm{S}=0}^{\infty} W_{\mathrm{Q}, \mathrm{R}, \mathrm{~S}} H^{2 \mathrm{Q}+\mathrm{S}_{\rho} 2 \mathrm{R}+\mathrm{S}_{\cos } \mathrm{S}_{\phi} .}
\end{aligned}
$$

It is standard to simplify the exponents by making the index transformation

$$
u=2 Q+S, v=2 R+S, w=S
$$

Then the wavefront aberration expansion takes the form

$$
\mathrm{W}(\mathrm{H}, \rho, \phi)=\sum_{\mathbf{u}}^{\infty} \sum_{\mathbf{v}}^{\infty} \sum_{\mathbf{w}}^{\infty} \mathbf{W}_{\mathbf{u}, \mathrm{v}, \mathbf{w}} \mathrm{H}^{\mathbf{u}} \rho^{\mathbf{v}} \cos ^{\mathbf{w}} \phi .
$$

Care must be exercised with this form because the sums no longer run from 0 to infinity.

The order of a wavefront aberration term is defined as

$$
\text { Order }=2(Q+R+S)=u+v .
$$

The higher the order of the wavefront aberration, the higher spatial frequencies of the wavefront deformation it describes.

Radially symmetric systems contain only even order aberration terms. Table 11 lists the wavefront aberration terms through fourth order. A major objective of aberration theory is to calculate the coefficients. $W_{u, v, w}$ from the prescription for an optical system.

## The Lowest Order Terms

The zeroth and second order wavefront aberrations do not so much characterize image quality as describe where and with what relative phase the images occur. The terms more descriptive of image quality are the iourin order terms, known as the Seidel aberrations, and the higher order terms. They are not described here since the polarization aberrations will only be carried oui to second order.

## Constant Piston Error

The coefficient $W_{000}$ represents zero order piston error. The value of $W_{000}$ is the net change in optical phase between the object and image for the ray path along the optical axis. If

$$
L=\sum_{q} n_{q} \ell_{q}
$$

is the sum of optical path lengths down the optical axis, then

$$
W_{000}=\bmod (L, \lambda)
$$

is the remainder after dividing $L$ by the wavelength. Since the term, $W_{000}$ is independent of object and pupil coordinates, it represents a constant phase term.

This term is usuaily ignored in optical design because:
a) it is rarely important in noninterferometric applications,

## TABLE 11

## The Wavefront Aberrations

## Zero Order

$W_{000} \quad 1$
Piston Error

| $W_{200}$ | $H^{2}$ |
| :--- | :--- |
| $W_{112}$ | $H \rho \cos \phi$ |
| $W_{020}$ | $\rho^{2}$ |

Fourth Order. Seidel Aberrations

Piston Error
Tilt
Defocus

| $\mathbf{W}_{400}$ | $\mathbf{H}^{4}$ |
| :--- | :--- |
| $\mathbf{W}_{11}$ | $\mathbf{H}^{3} \rho \cos \phi$ |
| $W_{220}$ | $\mathbf{H}^{2} \rho^{2}$ |
| $W_{222}$ | $\mathbf{H}^{2} \rho^{2} \cos ^{2} \phi$ |
| $W_{131}$ | $\mathbf{H}^{3} \rho^{3} \cos \phi$ |
| $W_{040}$ | $\boldsymbol{\rho}^{4}$ |

Piston Error
Distortion
Field Curvature
Astigmatism
Coma
Spherical Aberraticn
b) its exact value in optical hardware is easily several wavelengths away from the design value due to tolerances,
c) it can be readily adjusted with little perturbations to the other aberrations.

## Defocus

The $W_{020} \rho^{2}$ term is defocus. The functional form, $\rho^{2}$, is the paraxial equation for a spherical wave. Defocus describes the difference between the radius of curvature of the transmitted wavefront and the reference wavefront. Thus it represents a shift of focus out of the Gaussian image plane. Since the defocus term is independent of field coordinate, it describes a uniform focus shift to a plane parallel to the paraxial image plane. Higher order terms of the form $\mathrm{H}^{\mathrm{Q}}{ }^{\rho^{2}}$; such as $W_{220} \mathrm{H}^{2} \rho^{2}$ and $W_{420} \mathrm{H}^{4} \rho^{2}$, characterize nonuniform focal shifts or field curvature.

Tilt
The term $W_{111} \mathrm{H} \rho \cos \phi$ is tilt. Tilt is a linear change to the wavefront, equivalent to a pivot of the wavefront about the x axis. The tilt of the wavefront is zero for $\mathrm{H}=0$ and increases linearly as the object point moves off axis. This tilt moves the image without changing its form, exactly as a linear shift displaces a Fourier transform. Tilt produces a magnification change, since the image is uniformly enlarged.

## Quadratic Pistcn Error

The term $\mathrm{W}_{200} \mathrm{H}^{2}$ is quadratic piston error. It is the quadratic portion of the variation of the optical path length of the chief ray as a function of object coordinate. Since it is an "absolute phase" term which doesn't affect image quality. it is usually safely ignored.

## Jones Matrix Form for the Wavefront Aberrations

In aberration theory no mention is usually made of the polarization of the incident and transmitted wavefronts. This is the default assumption of geometrical optics, the optical system is assumed to transnit a beam of uniform intensity and uniform polarization state equal to the incident amplitude and polarization state. In the absence of further information, this is the most logical and safest assumption.

Assume that the system is nonpolarizing; then the transmitted wavefront has the same shape for all incident polarization states. Additionally assume that the transmitted wavefront has a uniform amplitude for all incident polarizations and for all ray paths. Then the Jones matrix and $\mathbf{C}$ vector as a function of system coordinates can immediateiy be witten. The Jones matrix and $C$ vector for a nonpolarizing uniform phase shift of W waves at constant amplitude are:

$$
\mathbf{J}=e^{j 2 \pi W / \lambda}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right], \text { and, } \quad \vec{C}=e^{j 2 \pi W / \lambda}[1,0,0,0] .
$$

Thus the Jones matrix and $C$ vector expansions for a uniformly transmitting nonpolarizing system described by the wave aberration expansion are:

$$
\begin{aligned}
& \mathbf{J}(H, \rho, \phi)=\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] \exp \left[\frac{j 2 \pi}{\lambda} \sum_{\mathbf{u}} \cdot \sum_{\mathbf{v}} \sum_{\mathbf{w}} W_{u, v, w} H_{\rho}^{\mathbf{u}^{v} \mathbf{v o s}^{w} \phi}\right] .
\end{aligned}
$$

Here the wavefront polynomial expansion has been used to describe the variation of the phase of the complex coefficient $c_{0}$ of the $\sigma_{0}$ "identity" polarization matrix. These equations embody the default polarization assumption in geometrical optics.

## CHAPTER 6

## POLARIZATION ABERRATIONS

## Introduction

This chapter presents a Taylor series description of the instrumental polarization associated with paraxial rays through radially symmetric systems of homogeneous and isotropic optical elements. The results are obtained in a form very similar to the wavefront aberrations. In particular, terms closely related to defocus, tilt, and piston error as well as the Seidel and higher order aberrations can be associated with all eight of the basis Jones matrices. Since polarization and retardance effects are typically orders of magnitude smaller than wavefront effects, fewer terms are needed for a sufficient description. A msthod of calculating aberration coefficients for specific systems is developed in Chapter 7.

## Prior Work

Many authors have pointed out the effects of, and the instrumental limitations due to, instrumental polarization in optical systems but few systematic analyses have been putlished. Papers which discuss instrumental polarization effects in spectrometers include Baur, Lewis and Hull (1980), Breckenridge (1971). Clout and Heddle (1969). Stewart and Gallaway (1962), and Woolsey and McConkey (1968). Howell (1979) performs a Mueller matrix analysis of the instrumental polarization of a radiometer. Stenflo (1978) discusses the limitations imposed on solar magnetic field measurements due to instrumental polarization. Instrumental polarization effects in phased array optical systems are discussed by Hege. Beckers. Strittmatter and

McCarthy (1985), McCarthy, Strittmatter, Hege and Low (1982), Shellan (1985), and West (1985).

One of the few systematic analyses of polarization aberration effects is due to Kubota and Inoue (1959) who analyzed image formation in polarizing microscopes. In their interference microscope, light traverses a linear polarizer, a microscope objective, reflects or scatters from a sample, travels back through the microscope objective and then passes through a polarization analyzer. The microscope is measuring polarization changes due to the sample. The rotation of the plane of linear polarization for light rays at nonnormal incidence at the lens surfaces reduces the sensitivity of the microscope. The authors present a detailed analysis of this polarization rotation arising at the lens surfaces. They also analyze the distribution of polarized light in the exit pupil and calculate the point spread function due to linear polarization defocus, a polarization aberration term introduced later in this chapter. They show that the point spread function for the light remaining in the incident polarization state loses its radial symmetry becoming bilaterially symmetric. The point spread function for the light coupled into the orthogonal polarization state has a four-fold symmetry and resembles the diffraction pattern in the presence of astigmatism.

Inoue and Hyde (1957) designed a polarization compensator for interference microscopes to compensate for the linear polarization defocus introduced by fast microscope objectives which works as follows. After the light has double passed the microscope objective, it passes through the polarization compensator, a half wave linear retarder followed by a short radius meniscus lens. The retarder is aligned with the initial polarizer. Linearly polarized light which is aligned with the retarder axes is unaffected by the retarder. Light whose linear polarization is at an angle $\delta$
with respect to the retarder axis is transmitted at an angle $-\delta$; the major axis is flipped about the retarder axis. The meniscus lens is chosen to introduce polarization rotation equal to that introduced by double-passing the microscope objective. The half wave retarder and meniscus lens effectively cancel the on-axis polarization aberration of the microscope objective. This polarization compensator greatly improves the quality of images from interference microscope by reducing the instrumental polarization and has been used with success in many microscopes.

Knowlden (1981) analyzed the instrumental polarization due to coating effects on two optical systems: an $\mathrm{f} / 1$ parabola used in collimated light, and a reflecting conical optical element with a $45^{\circ}$ half angle. He wrote a computer program to evaluate angle of incidence effects due to the coatings. Several plots of the variztion of amplitude and phase across the exit pupil for the different polarization components clearly show the effects of linear polarization defocus and linear retardance defocus. No analysis was done off axis.

Fainman and Shamir (1984) analyzed the diffraction of polarized light across an interface. Using a fundamentally different approach than is taken here (diffraction as opposed to ray-based), they derive the on-axis instrumental polarization effects at a spherical interface due to linear polarization defocus. Photographs of this effect are included.

Chipman (1985), using the Mueller calculus, performed an analysis of the second order polarization aberrations of uncoated lenses, treating the instrumental polarization arising from the Fresnel equations. Equations were provided to calculate aberration coefficients for sequences of uncoated lenses.

## The Polarization Aberration Expansion

The wavefront polynomial expansion describes the variation of the optical path difference through an optical system as a function of ray coordinates. A closely related expansion is presented for all four basis polarization matrices. This polarization aberration expansion for radially symmetric systems uses a very similar polynomial expansion to describe all eight basis polarization vectors. The principal difference is a modified form for the linear polarization and linear retardance terms since these involve both a magnitude and an orientation.

The Fresnel equations and other common weak polarizer expressions are relatively "weak" functions of the angle of incidence when compared with Snell's law. Thus it is possible to describe polarization variations using a lower order expansion than is necessary in expanding Snell's law for the wavefront aberrations.

## Amplitude, Phase and Complex Terms

It has been shown how the eight forms of polarization behavior can be characterized by four complex numbers, for example, the four elements of either the Jones matrix or the C vector. Here, complex polarization aberration coefficients $\mathrm{P}_{\mathrm{k}, \mathrm{u}, \mathrm{v}, \mathrm{w}}$ are used. It should be emphasized that the amplitude and phase of the coefficients are generally unrelated referring to different aspects of the instrumental polarization. The amplitude part of the coefficient A describes amplitude and polarization effects while the phase part $\Phi$ describes phase and retardance. The amplitude and phase of the polarization aberration coefficient P and the subscripts k , $u, v$, and $w$ are defined as follows:

$$
P_{k, u, v, w}=A_{k, u, v, w} e^{j \Phi_{k, u}, v, w}
$$

where: $k$ is the type of polarization behavior,
u is the order of the H dependence,
v is the order of the $\rho$ dependence, and,
$w$ is the order of the $\phi$ dependence.

The indices u,v and $w$ are used exactly as they were for the wavefront aberrations.

## Polarization Aberration Expansion

The polarization aberration expansion of the Jones matrix for a rotationally symmetric system is

$$
J(H, \rho, \phi)=\sum_{k=0}^{3} c_{k}(H, \rho, \phi) \sigma_{k}
$$

The $\mathbf{C}$ vector coefficients in this polarization aberration expansion are:

$$
\begin{aligned}
\mathrm{c}_{0}(\mathrm{H}, \rho, \phi)= & \mathrm{A}_{0000}+\hat{A}_{0200} \mathrm{H}^{2}+\hat{A}_{0111} \mathrm{Hi} \rho \cos \phi+\mathrm{A}_{0020} \rho^{2}+ \\
& \mathrm{j}\left(\Phi_{0000}+\Phi_{0200} \mathrm{H}^{2}+\Phi_{0111} \mathrm{H} \rho \cos \phi+\Phi_{0020} \rho^{2}\right) . \\
\mathrm{c}_{1}(\mathrm{H}, \rho, \phi)= & \mathrm{A}_{1000}+\mathrm{A}_{1200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\mathrm{~A}_{1111} \cos \phi-\mathrm{A}_{2111} \sin \phi\right)+\rho^{2}\left(\mathrm{~A}_{1022} \cos 2 \phi-\mathrm{A}_{2022} \sin 2 \phi+\right. \\
& \mathrm{j}\left(\Phi_{1000}+\Phi_{1200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\Phi_{1111} \cos \phi-\Phi_{2111} \sin \phi\right)+\rho^{2}\left(\Phi_{1022} \cos 2 \phi-\Phi_{2022} \sin 2 \phi\right)\right), \\
\mathrm{c}_{2}(\mathrm{H}, \rho, \phi)= & \left.\mathrm{A}_{2000}+\mathrm{A}_{2200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\mathrm{~A}_{2111} \cos \phi+\mathrm{A}_{1111} \sin \phi\right)+\rho^{2}\left(\mathrm{~A}_{2022} \cos 2 \phi+\mathrm{A}_{1022} \sin 2 \phi\right)\right) \\
& j\left(\Phi_{2000}+\Phi_{2200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\Phi_{2111} \cos \phi+\Phi_{1111} \sin \phi\right)+\rho^{2}\left(\Phi_{2022} \cos 2 \phi+\Phi_{1022} \sin 2 \phi\right)\right) . \\
\mathrm{c}_{3}(\mathrm{H}, \rho, \phi)= & \left.\mathrm{A}_{3000}+\mathrm{A}_{3200} \mathrm{H}^{2}+\mathrm{A}_{3111} \mathrm{H} \rho \cos \phi+\mathrm{A}_{3020} \rho^{2}\right]+ \\
& j\left(\Phi_{3000}+\Phi_{3200} \mathrm{H}^{2}+\Phi_{3111} \mathrm{H} \rho \cos \phi+\Phi_{3020} \rho^{2}\right) .
\end{aligned}
$$

## Description of Terms

This polarization aberration expansion for a radially symmetric system has thirty two terms to second order, which arise from four terms in each of the eight degrees of freedom of the Jones matrix. The terms are grouped as follows:

| $\mathbf{A}_{0, u, v, w}$ | Amplitude terms |
| :--- | :--- |
| $\mathbf{A}_{1, u, v, w}$ | Linear polarization terms |
| $\mathbf{A}_{2, u, v, w}$ | Diagonal polarization terms |
| $\mathbf{A}_{3, u, v, w}$ | Circular polarization terms |
| $\Phi_{0, u, v, w}$ | Wavefront or phase terms |
| $\Phi_{1, u}, \mathbf{v , w}$ | Linear retardance terms |
| $\Phi_{2, u, v, w}$ | Diagonal retardance terms |
| $\Phi_{3, u, v, w}$ | Circular retardance terms |
| $\mathbf{P}_{k, 0,0,0}$ |  |
| $\mathbf{P}_{k, 2,0,0}$ | "Constant Piston" terms |
| $\mathbf{P}_{k, 1,1,1}$ | "Quadratic Piston" terms |
| $\mathbf{P}_{k, 0,0,0}$ | "Tilt" terms |
| $\mathbf{P}_{k, 0,2,2}$ | "Scalar Defocus" terms |
|  | "Vector Defocus" terms |

The names of the wavefront aberrations: piston, quadratic piston, defocus and tilt, are used in an extended sense to describe variations of components of the Jones vector which share the same functional dependence as the wavefront aberrations. Defocus refers to a $\rho^{2}$ variation. Tilit reiers to a $\mathrm{H} p \cos \phi$ variation. Constant piston refers to a constant function. Quadratic piston refers to an $H^{2}$ variation. So "circular retardance tilt" is the $\mathrm{H} \rho \cos \phi$ variation of circular retardance.

This polarization aberration expansion is an equation which spans all possible second order variations of the Jones matrix, just as the second order wavefront aberration expansion spans the set of all second order wavefront variations. This polarization aberration expansion characterizes quadratic variations of all forms of wavefront, amplitude, polarization and retardance.

This polarization aberration expansion is a summation of terms in the different Pauli spin matrix components, not a product. Thus the four C vector elements should be pictured as acting in parallel, almost side by side in the aperture, but not in series. Each term describes an amount of a particular form of polarization, independent of the other contributions.

An "aberration term" is to be considered as containing all the algebraic terms in the expansion with the same coefficient. Most of the coefficients occur only once and the aberration term contains oniy one algebraic term. The exceptions are the terms, $A_{1111}, \Phi_{1111}, A_{1022}, \Phi_{1022}, A_{2111}, \Phi_{2111}, A_{2022}, \Phi_{2022}$. These aberration terms have components both along the axes and at 45 degrees.

## Single Surface Aberrations

in this section the zero and second order polarization aberration coefficients for a paraxial spherical wave incident at a single homogeneous interface are derived.

Consider a single weakly polarizing spherical surface illuminated by a spherical optical beam. Using the notation of Chapter 3 for weak polarizers, the second order C vector in s-p coordinates for a surface as a function of angle of incidence $i$ is

$$
\vec{Z}(i)=\left(d_{09}, d_{10}, d_{20}, d_{30}\right)+i^{2}\left(d_{02}, d_{12}, d_{22}, d_{32}\right) .
$$

The d's are complex numbers which incorporate both polarization and retardance effects. Terms $d_{20}, d_{22}, d_{30}$ and $d_{32}$ describe diagonal and circular polarization. These terms are zero for isotropic interfaces but are included here for completeness, For an arbitrary plane of incidence, $\theta$, the $\mathbf{C}$ vector is

$$
\begin{aligned}
\overleftrightarrow{d}(\mathrm{i}, \theta) & =\left(d_{00}, d_{10} \cos 2 \theta-d_{20} \sin 2 \theta, d_{20} \cos 2 \theta+d_{10} \sin 2 \theta, d_{30}\right) \\
& +i^{2}\left(d_{02}, d_{12} \cos 2 \theta-d_{22} \sin 2 \theta, d_{22} \cos 2 \theta+d_{12} \sin 2 \theta, d_{32}\right) .
\end{aligned}
$$

Let the angles of incidence for the marginal and chief rays be $\mathrm{i}_{\mathrm{m}}$ and $\mathrm{i}_{\mathrm{c}}$. The angle of incidence for a given ray in terms of its object and entrance pupil coordinates is

$$
\mathrm{i}^{2}=\mathrm{H}^{2} i_{\mathrm{c}}^{2}+2 H \rho \cos \phi \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \mathrm{i}_{\mathrm{m}}^{2} .
$$

Expressions for the $\mathbf{C}$ vector elements $\mathrm{c}_{0}(\mathrm{H}, \rho, \phi)$ and $\mathrm{c}_{3}(\mathrm{H}, \rho, \phi)$ are detrrmined by substitution of $\mathbf{i}^{\mathbf{2}}$ into ${ }^{\text {U }}(\mathrm{i}, \boldsymbol{\theta})$ as:

$$
\begin{aligned}
& c_{0}(H, \rho, \phi)=d_{00}+d_{02}\left[H^{2} i_{c}^{2}+2 H \rho \cos \phi i_{c^{2}} i_{m}+\rho^{2} i_{m}^{2}\right] . \\
& c_{3}(H, \rho, \bar{\phi})=d_{30}+d_{32}\left[H^{2} i_{c}^{2}+2 H \rho \cos \phi i_{c} i_{m}+\rho^{2} i_{m}^{2}\right] .
\end{aligned}
$$

The $C$ vector elements $c_{1}(H, \rho, \phi)$ and $c_{2}(H, \rho, \phi)$ ase ohtained by substituting the paraxial expressions for $\mathrm{i}^{2}, \sin 2 \theta$ and $\cos 2 \theta$. These elements are:

$$
\begin{aligned}
& \mathrm{c}_{1}(\mathrm{H}, \rho, \phi)=\mathrm{d}_{10}+\mathrm{d}_{12}\left[\mathbf{H}^{2} \mathrm{i}_{\mathrm{c}}^{2}+2 H \rho \cos \phi \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \cos 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right] \\
& +\mathrm{d}_{22}\left[2 \mathrm{H} \rho \sin \phi \mathrm{i}_{\mathrm{c}_{\mathrm{i}}}+\rho^{2} \sin 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right] . \\
& \mathrm{c}_{2}(\mathrm{H}, \rho, \phi)=\mathrm{d}_{2 \underline{n}}+\mathrm{d}_{22}\left[\mathrm{H}^{2} \mathrm{i}_{\hat{\varepsilon}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\overline{\mathrm{I}}}+\rho^{2} \cos 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right] \\
& +\mathrm{d}_{12}\left[2 \mathrm{H} \rho \sin \phi \mathrm{i}_{\mathrm{c}_{\mathrm{i}}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \sin 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right] .
\end{aligned}
$$

Equating the $c_{k}(H, \rho, \phi)$ 's with the polarization aberration expansion yields the following single surface polarization aberration contributions:

$$
\begin{aligned}
& P_{k, 0,0,0}=d_{k, 0} \cdot \\
& P_{k, 2,0,0}=d_{k, 2} i_{c}^{2} \\
& P_{k, 0,2,0}=d_{k, 2} i_{m}^{2} \\
& P_{k, 1,1,1}=2 d_{k, 2} i_{c} i_{m} .
\end{aligned}
$$

The second order single surface polarization aberration coefficients simply contain the angular dependence of the polarization times the appropriate functions of angle of incidence.

## Discussion of the Terms

In this section the form of the particular polarization aberrations terms is described. A distinction is made between scalar and vector aberrations. The
wavefront aberrations are scalar aberrations, single valued functions of object and pupil coordinates. The linear polarization and linear retardance aberrations are vector aberrations since a magnitude and orientation is associated with these at each point. Amplitude, circular polarization and circular retardance aberrations are scalar since they are single valued and range positive and negative.

## The Geometric Origin of Tilt and Piston

Figure 12 (bottom) shows the chief and limiting rays at an interface for objects on axis and at the edge of the field of view. Figure 12 (top) is a plot of the value of the angle of incidence along the $y$ axis as a function of $\rho$. Tilt terms naturally occur because as the object point moves off axis, the angle of incidence increases at one edge of the beam and decreases at the other edge. Tilt contains the first order portion of this correction. Figure 13 shows the off-axis angle of incidence squared and the decomposition of this into defocus, tilt and piston terms. These terms are required to describe a quadratic variation whose vertex is located at an arbitrary position on the $y$ axis because

$$
x^{2}+(y-a)^{2}=x^{2}+y^{2}-2 y a+a^{2} .
$$

Since $a$ is a linear function of $H, a=k H$, the quadratic polarization variation takes the functional form

$$
\begin{aligned}
c\left(x^{2}+(y-k H)^{2}\right) & =c\left(x^{2}+y^{2}\right)-2 c y k H+\mathrm{ck}^{2} \mathrm{H}^{2} \\
= & \mathrm{P}_{020} \rho^{2}+\mathrm{P}_{111} 2 \mathrm{H} \rho \cos \phi+\mathrm{P}_{200} \mathrm{H}^{2} .
\end{aligned}
$$

$\mathrm{P}_{020}, \mathrm{P}_{111}$, and $\mathrm{P}_{200}$ are the defocus, tilt and quadratic piston aberration coefficients. Tilt and piston terms arise naturally from decentered defocus. Similarly, the fourth order wavefront aberrations coma, astigmatism, field curvature and distortion arise


Figure 12 On and Off-Axis Angle of Incidence
The angle of incidence in the meridional plane is shown as a function of pupil coordinate both on and off axis.


Figure 13 The Angle of Incidence Squared
The angle of incidence squared in the meridional plane is depicted for an off-axis object and its decomposition into defocus. tilt and piston.
naturally from decentered spherical aberration,

$$
\begin{aligned}
\mathrm{c}\left(\rho-\mathrm{H}_{0}\right)^{4}= & \mathrm{W}_{040} \rho^{4}+\mathrm{W}_{131} \mathrm{H}^{3} \cos \phi+\mathrm{W}_{220} \mathrm{H}^{2} \rho^{2}+W_{222} \mathrm{H}^{2} \rho^{2} \cos ^{2} \phi \\
& +\mathrm{W}_{311} \mathrm{H}^{3} \rho \cos \phi+\mathrm{W}_{400} \mathrm{H}^{4} .
\end{aligned}
$$

## Scalar Polarization Aberrations

The four sets of scalar polarization aberrations are: amplitude, phase (or wavefront), circular polarization and circular retardance. All are strict functional analogues of the wavefront aberrations except that they describe other aspects of the Jones matrix than the phase of the identity matrix.

Figure 14 shows contour plots of the scalar aberrations, tilt, defocus and piston. Position in the graph represents position in the pupil.

Figure 15 shows representations of the circular aberrations, defocus, tilt and piston. The representation is the same for either circular polarization aberrations or circular retardance aberrations, the difference being whether a circle represents circular polarization or circular retardance. The size of the circle is the magnitude of the polarization or retardance. The arrow distinguishes between left and right circular polarization. Both positive and negative values of the aberrations are shown.

## Scalar Constant Piston Terms

The scalar constant piston terms are:

| Amplitude Constant Piston, | $\mathbf{A}_{0000} \sigma_{0}$, |
| :--- | :--- |
| Phase Constant Piston, | $\Phi_{0000} \sigma_{0}$, |
| Circular Polarization Constant Piston, | $\mathbf{A}_{3000} \sigma_{3}$, |
| Circular Retardance Constant Piston, | $\Phi_{5000} \sigma_{3}$. |



Piston

Figure 14 The Scalar Aberrations, Defocus. Tilt and Piston


Figure 15 Circular Polarization Defocus, Tilt and Piston
The variation of circular polarization with pupii for positive and negative amounts of circular polarization defocus, tilt and piston.

Scalar constant piston characterizes the normal incidence polarization of a surface or the "down-the-axis" instrumental polarization of a system. This polarization is a uniform polarization "bias" present at all pupil positions for all object coordinates. As Figure 14 and 15 show, scalar constant piston is constant inside the pupil. $\mathbf{A}_{0000}$ is the net amplitude transmittance down the optical axis while $\Phi_{0000}$ is the net phase.

## Scalar Defocus Terms

The scalar defocus terms are:

| Amplitude Defocus, | $\mathbf{A}_{0020} \rho^{2} \sigma_{0}$, |
| :--- | :--- |
| Phase Defocus, | $\Phi_{0020} \rho^{2} j \sigma_{0}$. |
| Circular Polarization Defceus, | $\mathbf{A}_{3020} \rho^{2} \sigma_{3}$. |
| Circular Retardance Defocus, | $\Phi_{3020} \rho^{2} j \sigma_{3}$. |

The aberration terms $\mathrm{A}_{0020}, \Phi_{0020}, \mathbf{A}_{3020}$ and $\Phi_{\text {s020 }}$ describe the defocus-like ( $\rho^{2}$ ) variation of instrumental polarization in the exit pupil present for on-axis object points. This term is present with the same magnitude for off-axis objects as for onaxis objects.

The wavefront defocus term $\Phi_{0020}$ is the "usual" wavefront defocus ( $\mathrm{W}_{020}$ ) described in the wavefront aberration section. If $\Phi_{0020}$ is greater than zero, the defocused wavefront leads the Gaussian reference surface. If it is less than zero it follows the Gaussian reference surface.

The term $\mathbf{A}_{0020}$ is amplitude defocus. It describes the amplitude of the transmitted wavefront, not its shape. When amplitude defocus is present there is a quadratic apodization of the pupil for ail object points. This apodization is due to the optical system, not to intensity variations in the incident light as in the case of
the Gaussian profile of a laser beam. For negative $\mathrm{A}_{0020}$, the center of the pupil is brighter and the pupil becomes dimmer quadratically with pupil radius. For positive $A_{0020}$ the edge of the pupil is brighter. Figure 16 shows one way that amplitude defocus occurs. This figure shows light propagating through a typical convex lens with spherical surfaces. To the side is a plot of the path length through the center region of the lens as a function of pupil radius for an on-axis object. The path length through the lens is a constant minus a quadratic term. If the lens is fabricated from weakly absorbing neutral density filter giass, the transmitted light has a radial amplitude distribution of the form

$$
A(\rho)=A_{0000}+A_{0020} \rho^{2}
$$

as is shown in Figure 17. Amplitude defocus also arises from angle of incidence dependent reflection or transmission variations.

The term $\Phi_{3020}$ is circular retardance defocus, the quadratic variation of circular retardance with pupil radius. Referring again to Figure 16, now let this lens formed from two thin glass shells filled with sugar water or any other suitable circularly birefringent material. The net circular atardance for a ray is proportional to the length of the ray path through the circularly birefringent medium, which is proportional to $\rho^{2}$ for the on-axis object point. This lens produces a radially symmetric quadratic variation of circular retardance. For left and right circularly polarized incident light, the amount of defocus due to the circular birefringence is equal and opposite. The focal length of the lens is different for left and right circular polarized light. For linearly polarized light, the polarization rotation angle is maximum for the center ray, and decreases quadratically toward the edge of the pupil.


Figure 16 Optical Path Length through a Lens on axis as a function of pupil coordinate.

## Transmitted Amplitude



## Pupil Coordinate

Figure 17 Amplitude Transmittance of Weakly Absorbing Lens as a Function of Pupil Coordinate

Circular polarization defocus, $A_{3020}$ is similar to $\Phi_{3020}$ except that it is an amplitude rather than phase effect, the pupil possessing different quadratic apodizations which are circular polarization dependent.

## Scalar Tilt Terms

The scalar tilt terms are:

| Amplitude Tilt, | $\mathrm{A}_{0111} \mathrm{H} \rho \cos \phi \sigma_{0}$, |
| :--- | :--- |
| Phase Tilt. | $\Phi_{0111} \mathrm{H} \rho \cos \phi \mathrm{j} \sigma_{0}$, |
| Circular Polarization Tilt. | $\mathbf{A}_{3111} \mathrm{H} \rho \cos \phi \sigma_{3}$, |
| Circular Retardance Tilt. | $\Phi_{3111} \mathrm{H} \rho \cos \phi \mathrm{j} \sigma_{3}$. |

The four scalar tilt aberrations, $A_{011}, \Phi_{0111}$ (same as wavefront tilt, $W_{111}$ ), $\mathrm{A}_{311}$, and $\Phi_{311}$, are strictly analogous to tilt. The functional form of scalar tilt is $\mathrm{H} \rho \cos \phi=\mathrm{Hy}$.

The $\rho \cos \theta$ represents a linear variation of a parameter which is greater in the top (or bottom) of the pupil and less in the bottom (or top). This term is linear in the object coordinate. H , and is not present for on axis objects.

Wavefront tilt, $\Phi_{011}$, describes the tipping of the wavefront with respect to the Gaussian reference sphere. This displaces the image from the paraxial image location. Since the displacement is linear in H , tilt produces a magnification change in the system.

The presence of amplitude tilt, $\mathrm{A}_{0111}$, means that the top of the pupil is brighter than the bottom, or vice versa, for off-axis objects. This amplitude variation changes ihe structure of the diffraction pattern and also causes variations of fringe visibility in an interferometer.

The presence of circular retardance tilt, $\Phi_{3111}$, indicates more circular retardance in the top of the pupil and less in the bottom, or vice versa, for off-axis image points. If linear polarized light is incident on a system with only $\Phi_{3111}$, the plane of polarization of the light is rotated one way in the top of the pupil, undeviated across the middle of the pupil, and rotated the opposite direction across the bottom. If on the other hand right and left circular polarized light is incident on a system with pure circular retardance tilt, the transmitted wavefront has opposite tilts for the left and right circular polarizization states.

## Scalar Quadratic Piston Terms

The scalar quadratic piston terms are:

| Amplitude Quadratic Piston, | $\mathbf{A}_{0200} \mathrm{H}^{2} \sigma_{0}$, |
| :--- | :--- |
| Phase Quadratic Piston, | $\Phi_{0200} \mathrm{H}^{2} \mathrm{j} \sigma_{0}$, |
| Circular Polarization Quadratic Piston, | $\mathbf{A}_{\mathbf{3 2 0 0}} \mathrm{H}^{2} \sigma_{3}$, |
| Circular Retardance Quadratic Piston, | $\Phi_{3200} \mathrm{H}^{\mathbf{2} j \sigma_{5}}$, |

Quadratic piston terms describe a uniform change of a parameter across the pupil which occurs quadratically with the object coordinate. In a system with pure amplitude quadratic piston, $\mathbf{A}_{0200}$, the pupil is uniformly bright on-axis. As the object point moves off axis, the pupil brightness changes but the pupil remains uniformly illuminated. Amplitude quadratic piston describes a field dependent variation of the average amplitude transmittance of the system. Likewise, $\mathbf{A}_{3200}$ and $\Phi_{3200}$ describe object coordinate dependent variations of circular polarization and circular retardance which are uniform across the pupil.

## Vector Polarization Aberrations

The linear polarization and linear retardance aberrations are classed as vector aberrations since there is a magnitude and orientation associated with each aberration at each point. The vector aberrations have been further grouped into "linear" aberrations and "diagonal aberrations". Linear vector defocus is a purely radial or purely tangential function, depending upon the sign. Diagonal vector defocus is always at $\pm 45^{\circ}$ to radial. The other linear aberrations derive from an expansion of translated linear defccus. The diagonal aberrations derive from translated diagonal defocus.

Diagonal aberrations do not occur at homogeneous and isotropic radially symmetric interfaces, but do occur at anisotropic interfaces. They are included here to complete the full second order basis set of possible Jones matrix polarization aberrations.

Figure 18 shows the three second order linear aberrations as a function of pupil coordinates for both positive and negative aberration coefficients. The figures are the same for polarization or retardance aberrations, the difference being whether a line represents linear polarization or linear retardance. The location of a line represents a position in the pupil. The length of the line is the amount of linear polarization or retardance. The orientation of the line is the orientation of the associated linear polarization or linear retardance. Figure 19 is the corresponding plot for the three diagonal aberrations. Note that when the sign of a vector aberration changes, the associated polarizations rotate $90^{\circ}$.

## Vector Constant Piston Terms

The vector constant piston terms are:

Linear Polarization Constant Piston,
$\dot{A}_{1000} \sigma_{1}$.


Figure 18 The Second Order Linear Polarization Aberrations


# Diagonal Polarization Defocus 



## Diagonal Polarization Tilt



Diagonal Polarization Piston

Figure 19 The Second Order Diagonal Polarization Aberrations

| Linear Retardance Constant Piston, | $\Phi_{1000} \mathrm{j} \sigma_{2}$. |
| :--- | :--- |
| Diagonal Polarization Constant Piston, | $\mathbf{A}_{2000} \sigma_{2}$. |
| Diagonal Retardance Constant Piston, | $\Phi_{2000} \sigma_{2}$. |

The constant piston terms are independent of object and pupil coordinate. The vector constant piston terms describe the linear polarization and linear retardance down the axis of the optical system. The " 1000 " terms describe linear polarization components along the system $x-y$ axes while the " 2000 " terms describe components at $\pm 45^{\circ}$. For systems with pure " 1000 " and " 2000 ", the polarization in the exit pupil is a constant, independent of pupil coordinates $\rho$ and $\phi$ or object position $H$. This corresponds to the polarization aberration of an ideal linear polarizer or retarder whose polarization or retardance varies quadratically with field height.

## Vector Defocus Terms

The vector defocus terms are:

| Linear Polarization Defocus, | $\mathrm{A}_{1020} \rho^{2}\left(\sigma_{1} \cos 2 \phi-\sigma_{2} \sin 2 \phi\right)$, |
| :--- | :--- |
| Linear Retardance Defocus, | $\Phi_{1020} \rho^{2} j\left(\sigma_{1} \cos 2 \phi-\sigma_{2} \sin 2 \phi\right)$, |
| Diagonal Polarization Defocus, | $A_{2020} \rho^{2}\left(\sigma_{2} \cos 2 \phi+\sigma_{1} \sin 2 \phi\right)$, |
| Diagonal Retardance Defocus, | $\Phi_{2020} \rho^{2} j\left(\sigma_{2} \cos 2 \phi+\sigma_{1} \sin 2 \phi\right)$. |

Vector defocus describes the defocus-like variation of linear polarization and linear retardance which is quadratic in the pupil radius but constant in the objeci coordinate.

To understand the origin of the linear polarization defocus and linear retardance defocus refer to the plots of the paraxial angle of incidence in Figure 11. Since the Fresnel equations and other related amplitude transmission relations are
even functions of the angle of incidence, the lowest order of polarization variation present is proportional to the angle of incidence squared. Figure 20 is a plot showing, for an on-axis object point, the magnitude of the angle of incidence squared with the plane of incidence unchanged. This is the expected form of polarization variation at spherical interfaces with on-axis objects. If the interface is linearly polarizing, the lines represent weak linear polarization; if the weak polarizer is birefringent, the lines represent retardance. Figure 21 is a plot of linear polarization defocus showing the form for both positive and negative values of $\mathrm{A}_{1020}$. The positive form of linear polarization defocus occurs when tine $a_{s}>a_{p}$, as occurs with reflection from metals. Thus this aberration describes the on-axis linear polarization behavior of spherical metal mirrors. Conversely, the negative form of linear polarization defocus occurs when $a_{p}>a_{s}$, as occurs with refraction at an uncoated interface.

A detailed discussion of linear polarization defocus, although not by this name, is contained in Kubota and Incue (1959).

## Vector Tilt Terms

The vector tilt terms are:

Linear Polarization Tilt,
Linear Retardance Tilt,
Diagonal Polarization Tilt,
Diagonal Retardance Tilt.

$$
\begin{aligned}
& \mathrm{A}_{111} \mathrm{H} \rho\left(\sigma_{1} \cos \phi-\sigma_{2} \sin \phi\right), \\
& \Phi_{111} \mathrm{H} \rho \mathrm{j}\left(\sigma_{1} \cos \phi-\sigma_{2} \sin \phi\right), \\
& \mathrm{A}_{2111} \mathrm{H} \rho\left(\sigma_{2} \cos \phi+\sigma_{1} \sin \phi\right), \\
& \Phi_{2111} \mathrm{H} \rho \mathrm{j}\left(\sigma_{2} \cos \phi+\sigma_{1} \sin \phi\right) .
\end{aligned}
$$

Linear polarization tilt occurs for the same reason as scalar tilt. As the object moves off axis, the angle of incidence is reduced on one side of the aperture. and increased on the other side. So the angle of incidence dependent polarization


Figure 20 The Angle of Incidence Squared as a Function of Pupil Coordinate for an On-Axis Object



a. $P_{1020}$ greater than zero

b. $P_{1020}$ less than zero

Figure 21 Linear Defocus
and retardance are reduced on one side and increased on the other. The vector tilt terms have orthogonal polarization at diametrically oppnsite points in the pupil. Thus, the vector tilt aberrations add linear polarization or linear retardance to one side and subtract it from the other.

## Vector Quadratic Piston Terms

The vector quadratic piston terms are:

Linear Polarization Quadratic Piston, $\mathrm{A}_{1200} \mathrm{Fi}^{2} \sigma_{1}$,
Linear Retardance Quadratic Piston, $\Phi_{1200} \mathrm{H}^{2} \mathrm{j} \sigma_{1}$.
Diagonal Polarization Quadratic Piston, $\mathrm{A}_{2200} \mathrm{H}^{2} \sigma_{2}$,
Diagonal Retardance Quadratic Piston, $\boldsymbol{\Phi}_{2200} \mathrm{H}^{2} \mathbf{j} \sigma_{2}$.

The vector quadratic piston aberrations describe the average linear polarization and linear retardance across the pupil which occurs as the object moves off axis. This term does not average out when integrating over the pupil; it represents an overall polarization and retardance bias present for off-axis objects.

## CHAPTER 7

## CALCULATION OF ABERRATION COEFFICIENTS

## Introduction

$\hat{A}$ method is presented for calculating the second order polarization aberration coefficients of a radially symmetric optical system given the $\mathbf{C}$ vector power series for each interface. This method is limited to systems of homogeneous and isotropic interfaces such as lenses, mirrors and fine thin film coatings. The polarization associated with propagation through dichroic, birefringent or otherwise polarizing media is not treated here.

## Single Surface Aberrations for Amplitude Transmittance Relations

For homogeneous and isotropic interfaces characterized by amplitude transmittance relations, such as lenses, mirrors and ideal thin film coatings, the polarization aberrations at a interface simplify considerably. At these interfaces the Fresnel equations and related thin film equations are separable into $s$ and $p$ components, so the Jones matrices representing the interface in s-p coordinates are diagonal. The off-diagonal terms, diagonal polarization $\sigma_{2}$ and circular polarization $\sigma_{3}$, are not present. Further, with isotropic media, the $s$ and $p$ amplitude transmission coefficients at normal incidence must be equal. Thus the amplitude transmission functions for a coated or uncoated interface are expanded as:

$$
\begin{aligned}
a_{S}(i) & =\left(a_{0}+a_{S}, 2^{\left.i^{2}+\ldots\right)} e^{j\left(\delta_{0}+\delta_{2}, i^{\left.i^{2}+\ldots\right)}\right.}\right. \\
& =a_{0}\left(!+\left(A_{2}+a_{2}\right) i^{2}+\ldots\right) e^{j\left(\delta_{0}+\left(\Delta_{2}+\delta_{2}\right) i^{2}+\ldots\right)} \\
& \cong \operatorname{ri}\left(1+\left(A_{2}+j \Delta_{2}+a_{2}+j \delta_{2}\right) i^{2}+\ldots\right)
\end{aligned}
$$

$$
\begin{aligned}
a_{p}(i) & =\left(a_{0}+a_{p}, 2^{2}+\ldots\right) e^{j\left(\delta_{0}+\delta_{2}, p^{\left.i^{2}+\ldots\right)}\right.} \\
& =a_{0}\left(1+\left(A_{2}-a_{2}\right) i^{2}+\ldots\right) e^{j\left(\delta_{0}+\left(\Delta_{2}-\delta_{2}\right) i^{2}+\ldots\right)} \\
& \cong \tau\left(1+\left(A_{2}+j \Delta_{2}-a_{2}-j \delta_{2}\right) i^{2}+\ldots\right)
\end{aligned} .
$$

where:

$$
\begin{aligned}
& A_{2}=\frac{a_{\mathrm{s}, 2}+a_{\mathrm{p}, 2}}{2 a_{0}}, a_{2}=\frac{a_{\mathrm{s}, 2}-a_{\mathrm{p}, 2}}{2 a_{0}} \\
& \Delta_{2}=\frac{\delta_{\mathrm{s}, 2}+\delta_{\mathrm{p}, 2}}{2}, \delta_{2}=\frac{\delta_{\mathrm{s}, 2}-\delta_{\mathrm{p}, 2}}{2} .
\end{aligned}
$$

The normalized c vector coefficients are:

$$
d_{02}=A_{2}+j \Delta_{2} \text {, and, } d_{12}=a_{2}+j \delta_{2} .
$$

and,

$$
\tau=\mathrm{a}_{0} \mathrm{e}^{\mathrm{j} \delta_{0}}
$$

The s-p coordinate Jones matrix expansion to second order is

$$
J(i)=\left[\begin{array}{cc}
a_{s}(i) & 0 \\
0 & a_{p}(i)
\end{array}\right]=\tau\left(\sigma_{0}\left(1+\left(A_{2}+j \Delta_{2}\right) i^{2}+\sigma_{1}\left(a_{2}+j \hat{o}_{2}\right) i^{2}\right) .\right.
$$

The s-p coordinate $C$ vector expansion to second order is

$$
\left.\mathrm{U}_{\mathrm{i}}\right)=\tau(1,0,0,0)+\mathrm{i}^{2} \tau\left(\mathrm{~d}_{02}, \mathrm{~d}_{12}, 0,0\right) .
$$

The $x-y$ coordinate $C$ vector Taylor series for orientation of the plane of incidence $\theta$ is

$$
\overrightarrow{\mathrm{O}} \mathrm{i}, \theta)=\tau(1,0,0,0)+\mathrm{j}^{2} \tau\left(\mathrm{~d}_{02}, \mathrm{~d}_{12} \cos 2 \theta, \mathrm{~d}_{12} \sin 2 \theta, 0\right) .
$$

The normal-incidence polarization aberration terms (the constant piston terms) are zero:

$$
P_{1000}=P_{2000}=P_{3000}=0 .
$$

There is no polarization or retardance on axis, only the amplitude and phase transmission factor $\tau$.

All terms for the diagonal and circular polarization are zero:

$$
P_{2, u, v, w}=P_{3, u, v, w}=0
$$

Thus, the single surface $\mathbf{C}$ vector in paraxial coordinates is obtained by substituting $\mathrm{i}(\mathrm{H}, \rho, \phi)$ and $\theta(\mathrm{H}, \rho, \phi)$ into $\mathrm{Z}_{\mathrm{i}}(\mathrm{i}, \theta)$ yielding

$$
\begin{aligned}
& \mathrm{c}_{0}(\mathrm{H}, \rho, \phi)=\tau+\tau \mathrm{d}_{02}\left[\mathrm{H}_{\mathrm{c}}^{2} \mathrm{i}_{\mathrm{c}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \mathrm{i}_{\mathrm{m}}^{2}\right], \\
& \mathrm{c}_{1}(\mathrm{H}, \rho, \phi)=\tau \mathrm{d}_{12}\left[\mathrm{H}^{2} \mathrm{i}_{\mathrm{c}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}_{\mathrm{i}}}+\rho^{2} \cos 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right], \\
& \mathrm{c}_{2}(\mathrm{H}, \rho, \hat{\varphi})=\tau \mathrm{d}_{12}\left[2 \mathrm{H} \rho \sin \phi \mathrm{i}_{\mathrm{c}_{\mathrm{m}}}+\rho^{2} \sin 2 \phi \mathrm{i}_{\underline{m}}^{2}\right] . \\
& \mathrm{c}_{3}(\mathrm{H}, \rho, \phi)=0
\end{aligned}
$$

Since there is no diagonal polarization, the only contributions to $c_{2}$ arises from the rotation of linear polarization from the s-p coordinates into the $x-y$ coordinates.

## Example - Spherical Metal Mirror

As an example of the single surface polarization aberration expansion, the instrumental polarization function for reflention from a spherical mirror will be calculated. Figure 22 shows the cinfiguration and notation. Light from object point $P$ is incident on a spherical mirror and imaged at point $P^{\prime}$. The distance $s_{01}$ from $P$ to the vertex of the mirror V is positive when P is located to the left of the mirror and negative when $P$ is a virtual object. The radius of curvature $R$ of the mirror is negative for a concave mirror and positive for a convex mirror. The entrance pupil is located at the mirror and has semidiameter $r$. The height of the object at the edge of the field of view is $h$. Let $\left(0, y_{0}\right)$ and $\left(x_{1}, y_{1}\right)$ be the $x-y$ coordinates of a ray at the object and at the mirror. The paraxial normalized coordinates for this ray are:

$$
\begin{aligned}
& \rho=\frac{\sqrt{x_{1}^{2}+y_{1}^{2}}}{r}, 0<\rho<1, \\
& \theta=\arctan 2\left[-\frac{x_{1}}{y_{1}}\right],
\end{aligned}
$$



Figure 22 Spherical Mirror
R, Radius of curvature of mirror; $h$, Object height; $r$, Mirror semidiameter; $\mathrm{i}_{\mathrm{m}}$. Marginal ray angle of incidence; $i_{c}$. Chief ray angle of incidence.

$$
H=\frac{y_{0}}{h},-1<H<1 .
$$

The chief and marginal ray angles of incidence are:

$$
i_{c}=-\frac{h}{s_{01}}, \text { and, } \quad i_{m}=\frac{r}{s_{01}}-\frac{r}{R}
$$

As a numerical example, assume the mirror has an aluminum thin film overcoating with a complex index of refraction of

$$
\mathrm{n}=0.82-\mathrm{j} 5.99 .
$$

The C vector for reflection is

$$
\overrightarrow{\mathrm{C}}(\mathrm{i}, \theta)=\tau[1,0,0,0]+\mathrm{i}^{2}\left[0, \mathrm{~d}_{12} \cos 2 \theta, \mathrm{~d}_{12} \sin 2 \theta, 0\right] .
$$

Since $d_{02}$ is zero. the average amplitude reflectance is a constant. $\tau$. to second order as a function of angle of incidence. Due to $d_{12}, a_{s}$ and $a_{p}$ split symmetrically about $\tau$. The normal incidence amplitude reflectance $\tau$ is

$$
\tau=d_{00}=\frac{n-1}{n+1}
$$

The coefficient $d_{12}$ which characterizes the second order lintar polarization and linear retardance is

$$
d_{12}=-\frac{n-1}{n(n+1)} .
$$

For the aluminum example

$$
\tau=-0.907+\mathrm{j} 0.306 \text {, and, } \mathrm{d}_{12}=0.0705+0.142 \mathrm{j} .
$$

The normal incidence intensity reflectance $I_{0}$ is

$$
\mathrm{I}_{0}=\tau \tau^{*}=0.916
$$

The C vector for the mirror reflection. $\overrightarrow{\mathrm{C}}_{\mathrm{m}}$, is
$\vec{C}_{m}(\mathrm{H}, \rho, \phi)=\tau[1,0.0 .0]+\mathrm{d}_{12} \mathrm{H}^{2} \mathrm{i}_{\mathrm{c}}^{2}[0.1 .0 .0]-2 \mathrm{~d}_{12} \mathrm{H}_{\mathrm{i}} \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\mathrm{m}}[0 . \cos \phi . \sin \phi .0]$

$$
+d_{12} \rho^{2} i_{m}^{2}[0, \cos 2 \phi, \sin 2 \phi, 0]
$$

$$
\begin{aligned}
= & P_{0000}[1,0.0 .0]+P_{1200} \mathrm{H}^{2}[0.1 .0 .0] \\
& +P_{1111} \mathrm{H} p[0 . \cos \phi . \sin \phi .0]+P_{10200^{2}}[0 . \cos 2 \phi . \sin 2 \phi .0] .
\end{aligned}
$$

In terms of the system dimensions, the polarization aberration coefficients are:

$$
\begin{aligned}
& P_{0000}=\tau, \\
& P_{1200}=\frac{d_{12} h^{2}}{s_{01}^{2}}, \\
& P_{1111}=\frac{2 d_{12} h r}{s_{01}}\left[\frac{1}{s_{01}}-\frac{1}{R}\right] . \\
& P_{1020}=d_{12} r^{2}\left[\frac{1}{S_{01}}-\frac{1}{R}\right]^{2} .
\end{aligned}
$$

## Object at the Center of Curvature

Consider this mirror used for one-to-one imaging with the object at the center of curvature of the mirror and a magnification of minus one. Then

$$
s_{01}=R, P_{1111}=0 \text {, and, } P_{1020}=0
$$

For an on-axis object, $\mathrm{H}=0$, the mirror displays no polarization or retardance. This occurs because all rays from the center of curvature intersect the mirror at normal incidence. The only polarization term acting is the normal incidence amplitude transmission 7 . There is a uniform amplitude loss and a uniform phase shift but no pupil dependent polarization.

As the object print moves off axis from the center of curvature iit the paraxial approximation, the angle of incidence and plane of incidence are still equal for all rays from a given object point. The mirror acts as a uniform weak polarizer and uniform weak retarder for a given object point. in terms of the field of view, which is the chief ray angle of incidence $\mathrm{i}_{\mathrm{c}}$.

$$
P_{1200}=d_{12} i_{c} .
$$

The linear tilt and linear defocus are zero. There is field dependent polarization
and retardance but no pupil dependent variation.
For the aluminum coating example.

$$
P_{1200}=A_{1200}+j \Phi_{1200}=(0.0705+0.142 j) i_{c}=0.159 e^{j 1.08} i_{c}{ }^{2} .
$$

## Polarizat:on Aberrations for Cascaded Surfaces

The second order polarization aberration coefficients for a series of isotropic, weakly polarizing radially symmetric interfaces is derived.

## Paraxial Approximation

Since the polarization aberrations are being evaluated to second order in the angle of incidence. The difference between spheres, parabolas, conics or other radially symmetric aspherics does not occur at this order. The relevant shape parameter here is only the vertex radius of curvature. The angle and plane of incidence differences for these types of interfaces are the same at second order but differ at fourth order and higher.

## System Aberration Calculation

For surfaces $\mathrm{q}=1$ to $\mathbf{Q}$, each surface is characterized by three complex parameters from the normalized $C$ vector expansion: $d_{00}=\tau_{q} \cdot d_{0,2, q^{*}}$ and, $d_{1,2, q}$ The single surface polarization aberration coefficients are:

$$
\begin{aligned}
& \mathrm{P}_{0,0,0,0, q}={ }^{\top} \mathbf{q} \\
& \mathrm{P}_{0,2,0,0, q}=\mathrm{T}_{\mathrm{q}} \mathrm{~d}_{0,2, q} \mathrm{i}_{\mathrm{c}}^{2} \\
& P_{0,1,1,1, q}=2 \tau_{q}{ }^{d} 0,2, q i_{c} i_{m} \\
& P_{0,0,2,0, q}=\tau_{q}{ }^{d_{0,2, q}, i_{m}^{2}} \\
& P_{1,2,0,0, q}=\tau_{q} d_{1,2, q} i_{c}^{2} \\
& P_{1,1,1,!, q}=2 \tau_{q}{ }^{d}{ }_{1,2, q} i_{c} i_{m} \\
& P_{1,0,2,0, q}=\tau_{q} d_{i, 2, q} i_{m}^{2}
\end{aligned}
$$

The polarization aberration coefficients for the system are calculated by chain multiplying the single surface polarization aberration expressions and retaining terms to second order in H and $\rho$ where the order of a term is the sum of the powers of H and $\rho$,

$$
H^{u}{ }_{\rho} \mathbf{v}^{\cos }{ }^{\mathbf{w}} \phi, \text { order }=u+\mathbf{v} .
$$

The zero and second order Jones matrices for the $q^{\prime}$ th interface are:

$$
\begin{aligned}
\mathrm{J}_{\mathrm{q}}^{(0)}(\mathrm{H}, \rho, \phi) & =\mathrm{P}_{0,0,0,0, \mathrm{q}} \sigma_{0}=\mathrm{d}_{0,0, \mathrm{q}} \sigma_{0}=\tau_{\mathrm{q}} \sigma_{0} . \\
\mathrm{J}_{\mathrm{q}}^{(2)}(\mathrm{H}, \rho, \phi) & =\sigma_{0}\left[\mathrm{H}^{2} \mathrm{P}_{0,2,0,0, \mathrm{q}}+2 \mathrm{H} \rho \cos \phi \mathrm{P}_{0,1,1,1, \mathrm{q}}+\rho^{2} \mathrm{P}_{0,0,2,0, \mathrm{q}}\right] \\
& +g_{\mathrm{i}}\left[\mathrm{H}^{2} \mathrm{P}_{1,2,0,0, \mathrm{q}}+2 \mathrm{H} \rho \cos \phi \mathrm{P}_{1,1,1,1, \mathrm{q}}+\rho^{2} \cos 2 \phi \mathrm{P}_{1,0,2,0, \mathrm{q}}\right] \\
& +\sigma_{2}\left[2 \mathrm{H} \rho \sin \phi \mathrm{P}_{1,1,1,1, \mathrm{q}}+\rho^{2} \sin 2 \phi \mathrm{P}_{1,0,2,0, \mathrm{q}}\right] \\
& =\sigma_{\mathrm{q}} \tau_{\mathrm{q}} \mathrm{~d}_{02}\left[\mathrm{H}^{2} \mathrm{i}_{\mathrm{c}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}_{\mathrm{m}}}+\rho^{2} \mathrm{i}_{\mathrm{m}}^{2}\right] \\
& +\sigma_{1} \tau_{\mathrm{q}} \mathrm{~d}_{12}\left[\mathrm{H}^{2} \mathrm{i}_{\mathrm{c}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \cos 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right] \\
& +\sigma_{2} \tau_{q} \mathrm{~d}_{12}\left[2 \mathrm{H} \rho \sin \phi \mathrm{i}_{\mathrm{c}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \sin 2 \phi \mathrm{i}_{\mathrm{m}}^{2}\right] .
\end{aligned}
$$

Multiplication of the single surface Jones matrices yields

$$
J(H, \rho, \phi)=\prod_{q=Q,-1}^{1} J_{q}(H, \rho, \phi)=\prod_{q=Q,-1}^{1}\left[J_{q}^{(0)}+J_{q}^{(2)}(H, \rho, \phi)\right]
$$

Since $\mathrm{J}_{\mathrm{q}}^{(0)}$ is a constant function, independent of $\mathrm{H}, \rho$ and $\phi$, the ( $\mathrm{H}, \rho, \phi$ ) dependence can be dropped. This expression contains $2^{\mathbf{Q}}$ terms including one zero order term and Q second order terms.

The zero order Jones matrix is

$$
J^{(0)}=\prod_{q=Q,-1}^{1} J_{\mathbf{q}}^{(0)}(H, \rho, \phi)=\prod_{\mathbf{q}=1}^{Q} \tau_{\mathbf{q}}=\tau
$$

the system amplitude transmittance.
The second order Jones matrix is greatly simplified since, for isotropic surfaces, all zeroth order Jones matrices are a constant times the identity matrix $\sigma_{0}$.

The second order Jones matrix, which only includes products with a single second order term, is

$$
\mathrm{J}^{(2)}(\mathrm{H}, \mathrm{p}, \mathrm{\phi})=\mathrm{r} \sum_{\mathrm{q}=1}^{\mathrm{Q}} \mathrm{~J}_{\mathrm{q}}^{(2)}(\mathrm{H}, \mathrm{p}, \phi) .
$$

At second order the weakly polarizing isotropic interfaces do not display order dependence. The product of any two second order terms is fourth order. The order dependence enters at fourth and higher order. Second order is a simple sum of polarization contributions. Collecting the piston, tilt and defocus terms from the second order Jones matrix yields the coefficients for the system polarization aberration expansion to second order:

$$
\begin{aligned}
& P_{0,0,0,0}=\Upsilon, \\
& P_{0,2,0,0}=\Upsilon \sum_{q=1}^{Q} d_{0,2, q} i_{c}^{2}, \\
& P_{0,1,1,1}=2 \Upsilon \sum_{q=1}^{Q} d_{0,2, q} i_{c} i_{m} . \\
& P_{0,0,2,0}=\Upsilon \sum_{q=1}^{Q} d_{0,2, q} i_{m}^{2}, \\
& P_{1,2,0,0}=\Upsilon \sum_{q=1}^{Q} d_{1,2, q} i_{c}^{2} . \\
& P_{1,1,1,1}=2 \Upsilon \sum_{q=1}^{Q} d_{1,2, q} i_{c} i_{m} . \\
& P_{1,0,2,2}=\Upsilon \sum_{q=1}^{Q} d_{1,2, q} i_{m}^{2} .
\end{aligned}
$$

The other three zero order coefficients and the other six second order coefficients (diagonal and circular) are all zero:

$$
\begin{aligned}
& P_{1,0,0,0}=P_{2,0,0,0}=P_{3,0,0,0}=0 . \\
& P_{2,2,0,0}=P_{2,1,1,1}=P_{2,0,2,0}=0 . \\
& P_{3,2,0,0}=P_{3,1,1,1}=P_{3,0,2,0}=0 .
\end{aligned}
$$

The amplitude and polarization coefficients are the real parts of the $P$ coefficients

$$
A_{k, u, v, w}=\operatorname{Re}\left(P_{k, u, v, w}\right) .
$$

The phase and retardation coefficients are the imaginary parts

$$
\Phi_{\mathrm{k}, \mathrm{u}, \mathrm{v}, \mathrm{w}}=\operatorname{Im}\left(\mathrm{P}_{\mathrm{k}, \mathrm{u}, \mathrm{v}, \mathrm{w}}\right)
$$

The analysis is simplified when the calculation is performed between the entrance pupil and exit pupil. The entrance pupil $e$ is the spherical surface in object space centered on the object and conjugate with the system stop. Likewise. the exit pupii $e^{\prime}$ lies in "image space" and is the spherical surface centered on the image which is conjugate to the system stop. Each object and image point have their own entrance and exit pupil. Since $e$ and $e^{\prime}$ are conjugate (object and image to each other), the optical path length between $\mathbf{e}$ and $\mathbf{e}^{\prime}$ for all rays is constant to at least second order in the pupil coordinates so it can be assumed constant for this paraxial analysis. Performing the analysis between conjugate surfaces allows the quadratic wavefront aberration terms between surfaces to be discarded since they must all add to zero between the entrance and exit pupil.

## CHAPTER 8

## ALTERNATE POLARIZATION ABERRATION EXPANSIONS

## Introduction

The purpose of a polarization aberration expansion is to characterize variations of the instrumental polarization as a function of ray coordinates and wavelength. Since the instrumental polarization fuaction $\dot{j}(\overrightarrow{\mathrm{~h}}, \vec{\rho}, \lambda)$ is very difficult to obtain in closed form, useful and accurate approximations are sought. Any function $\mathbf{J}^{\prime}(\overrightarrow{\mathrm{h}}, \vec{j}, \lambda)$ which accurately approximates $\mathbf{J}(\overrightarrow{\mathrm{h}}, \overrightarrow{\boldsymbol{p}}, \lambda)$ over a range of $\overrightarrow{\mathrm{h}}, \overrightarrow{\boldsymbol{\gamma}}$, and $\lambda$ is a potentially useful polarization aberration function. The utility depends on the connections which can be made between the optical system prescription and the coefficients (or free parameters) in $\mathbf{J}^{\prime}(\overrightarrow{\mathrm{h}}, \vec{\rho}, \lambda)$. The expansion introduced in Chapter 6 is the "C vector expansion." It expresses $J^{\prime}(\vec{h}, \vec{j}, \lambda)$ as a quadratic function of $\vec{h}$ and $\vec{p}$. It is useful because these terms are easily related to the Taylor series representation of polarizing interfaces.

There are many other valid forms for second order polynomial aberration expansion of the instrumental polarization function. In this chapter, two closely related forms of the polarization aberration expansion are introduced. The first equation. "the exponential phase polarization expansion," places the phase and retardance comporents in exponential form but leaves the amplitude and retardance in polynomial form. The second equation, "the exponential polarization aberration expansion" places all components, both amplitude and phase, in exponential form. Both expansions reduce to the $C$ vector expansion in the weak polarizer limit and the
paraxial limit. These alternative forms offer advantages in handling strong polarizers and strong retarders. They have the disadvantage of additional mathematical complexity. These alternate expansions are only introduced here; a detailed analysis should appear in a later paper.

## The C Vector Aberration Expansion

The $\mathbf{C}$ vector form of the polarization aberration expansion introduced in Chapter 6 expresses the Jones vector for a rotationally symmetric optical system as

$$
J(H, \rho, \dot{\varphi})=\sum_{k=0}^{3} c_{\mathbf{k}}(H, \rho, \phi) \sigma_{k}
$$

The $C$ vector coefficients $c_{k}$ in this expansion are functions of the spatial variables in the object and pupil. The second order $C$ vector coefficients are:

$$
\begin{aligned}
c_{0}(H, \rho, \phi)= & A_{0000}+A_{0200} H^{2}+A_{0111} H \rho \cos \phi+A_{0020} \rho^{2} \\
& +j\left(\Phi_{0000}+\Phi_{0200} H^{2}+\Phi_{0111} H \rho \cos \phi+\Phi_{0020} \rho^{2}\right) \\
c_{1}(H, \rho, \phi)= & A_{1000}+A_{1200} H^{2}+H \rho\left(A_{1111} \cos \phi-A_{2111} \sin \phi\right)+\rho^{2}\left(A_{1022} \cos 2 \phi-A_{2022} \sin 2 \phi\right) \\
& +j\left(\Phi_{1000}+\Phi_{1200} H^{2}+H \rho\left(\Phi_{1111} \cos \phi-\Phi_{2111} \sin \phi\right)+\rho^{2}\left(\Phi_{1022} \cos 2 \phi-\Phi_{2022} \sin 2 \phi\right)\right. \\
c_{2}(H, \rho, \phi)= & A_{2000}+\dot{\beta}_{2200} H^{2}+H \rho\left(A_{2111} \cos \phi+A_{1111} \sin \phi\right)+\rho^{2}\left(A_{2022} \cos 2 \phi+A_{1022} \sin 2 \phi\right) \\
& +j\left(\Phi_{2000}+\Phi_{2200} H^{2}+H \rho\left(\Phi_{2111} \cos \phi+\Phi_{1111} \sin \phi\right)+\rho^{2}\left(\Phi_{2022} \cos 2 \phi+\Phi_{1022} \sin 2 \phi\right)\right)
\end{aligned}
$$

$$
\begin{aligned}
c_{3}(H, \rho, \phi)= & A_{3000}+A_{3200} H^{2}+A_{3111} H \rho \cos \phi+A_{3020} \rho^{2} \\
& +j\left(\Phi_{3000}+\Phi_{3200} H^{2}+\Phi_{3111} H \rho \cos \phi+\Phi_{3020} \rho^{2}\right)
\end{aligned}
$$

The $A$ 's and $\Phi$ 's are functions of the optical system configuration and the polarization properties of the optical elements.

This expansion breaks down for strong retardeis and large phases. The exact equation for an element with retardation $\delta(k=1,2,3)$ or phase $\delta(k=0)$ is (see Appendix
C)

$$
J=e^{j \delta \sigma_{k}}=\sigma_{0} \cos \delta+j \sigma_{k} \sin \delta
$$

Expanding J

$$
\mathrm{J} \cong \sigma_{0}+j \delta \sigma_{\mathrm{k}}-\frac{\delta^{2} \sigma_{0}}{2}-\frac{\mathrm{j} \delta^{3} \sigma_{k}}{6}+\ldots .(\delta<1 \text { radian })
$$

aand keeping terms to first order yields

$$
\mathbf{J} \cong \sigma_{0}+j \delta \sigma_{k} .(\delta \ll 1 \text { radian. })
$$

This is the order of accuracy of the C vector aberration expansion. This expression is accurate only for small values of $\delta$, either small phases or small retardances. This limitation is most acute in the phase, $\mathrm{k}=0$, which frequently needs to vary over many radians to describe an optical system. The two next higher order terms in the $\delta<1$ radian expression, describe: (second order) corrections to $\sigma_{0}$ to normaiize for the ïrst order $\sigma_{\mathrm{k}}$ term, and (third order) nonlinearities in $\sigma_{\mathrm{k}}$ as a function of $\delta$.

This limitation of the $\mathbf{C}$ vector aberration expansion to small phases can be sidestepped by factoring the phase out and handling it separately as a multiplicative factor. This works well since phase commutes with all the Pauli spin matrices. The next section applies this factoring approach to the phase and retardance.

## The Exponential Phase Polarization Aberration Expansion

When phase and retardance coefficients are placed in an exponent, the polarization aberration expansion will accurately add and subtract large values of phase or retardance. An example of this is the following equation, the unsymmetrized exponential phase polarization aberration expansion.

$$
J(H, \rho, \phi)=\left[\sum_{k=0}^{3} \rho_{k}(H, \rho, \phi) \sigma_{k}\right] \exp \left[\sum_{k=0}^{3} \theta_{k}(H, \rho, \phi) \sigma_{k}\right]
$$

## $=\left(\rho_{0} \sigma_{0}+\rho_{1} \sigma_{1}+\rho_{2} \sigma_{2}+\rho_{3} \sigma_{3}\right) \mathrm{e}$

$$
j\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)
$$

For a second order polynomial expansion of the spatial variables, the aberration coefficients are of the same form as for the C vector polarization abberation expansion:

$$
\begin{aligned}
& \rho_{0}(\mathrm{H}, \rho, \phi)=\mathrm{A}_{0000}+\mathrm{A}_{0200} \mathrm{H}^{2}+\mathrm{A}_{0111} \mathrm{H} \rho \cos \phi+\mathrm{A}_{0020} \rho^{2}, \\
& \Theta_{0}(\mathrm{H}, \rho, \phi)=\Phi_{0000}+\Phi_{0200} \mathrm{H}^{2}+\Phi_{0111} \mathrm{H}^{2} \cos \phi+\Phi_{0020} \rho^{2}, \\
& \rho_{1}(\mathrm{H}, \rho, \phi)=\mathrm{A}_{1000}+\mathrm{A}_{1200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\mathrm{~A}_{1111} \cos \phi-\mathrm{A}_{211} \sin \phi\right)+\rho^{2}\left(\mathrm{~A}_{1022} \cos 2 \phi-\mathrm{A}_{2022} \sin 2 \phi\right) . \\
& \theta_{1}(\mathrm{H}, \rho, \phi)=\Phi_{1000}+\Phi_{1200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\Phi_{1111} \cos \phi-\Phi_{2111} \sin \phi\right)+\rho^{2}\left(\Phi_{1022} \cos 2 \phi-\Phi_{2022} \sin 2 \phi\right) . \\
& \rho_{2}(\mathrm{H}, \rho, \phi)=\mathrm{A}_{2000}+\mathrm{A}_{2200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\mathrm{~A}_{2111} \cos \phi+\mathrm{A}_{1111} \sin \phi\right)+\rho^{2}\left(\mathrm{~A}_{2022} \cos 2 \phi+\mathrm{A}_{1022} \sin 2 \phi\right) . \\
& \Theta_{2}(\mathrm{H}, \rho, \phi)=\Phi_{2000}+\Phi_{2200} \mathrm{H}^{2}+\mathrm{H} \rho\left(\Phi_{2111} \cos \phi+\Phi_{1111} \sin \phi\right)+\rho^{2}\left(\Phi_{2022} \cos 2 \phi+\Phi_{1022} \sin 2 \phi\right) . \\
& \rho_{3}(\mathrm{H}, \rho, \phi)=\mathrm{A}_{3000}+\mathrm{A}_{3200} \mathrm{H}^{2}+\mathrm{A}_{3111} \mathrm{H} \rho \cos \phi+\mathrm{A}_{3020} \rho^{2} . \\
& \Theta_{3}(\mathrm{H}, \rho, \phi)=\Phi_{3000}+\Phi_{3200} \mathrm{H}^{2}+\Phi_{311} \mathrm{H} \rho \cos \phi+\Phi_{3020} \rho^{2} .
\end{aligned}
$$

In the weak polarizar limit, these A's anu' $\Phi$ 's have the same values as the $A$ 's and $\Phi$ 's in the C vector polarization aberration expansion.

One aesthetic aspect of the exponential form for the retarders is its order independent representation for the retardance terms. In expanding the exponential

$$
\begin{aligned}
\mathrm{e}^{\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)}= & \sigma_{0}+\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right) \\
& +\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right) \\
& +\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)+\ldots
\end{aligned}
$$

all order dependent terms in the expansion cancel. There are no nonzero contributions from products of the noncommuling terms, $\sigma_{1}, \sigma_{2}$ or $\sigma_{3}$. For every $\sigma_{1} \sigma_{2}=j \sigma_{3}$ there is a $\sigma_{2} \sigma_{1}=-j \sigma_{3}$ with the same coefficients. Noncommuting terms do not interact in this exponential; they act only on themselves and on $\sigma_{0}$, which commutes with everything. The coefficients, $\theta_{0}, \theta_{1}, \theta_{2}$, and $\theta_{3}$ represent pure costributions of phase or the three forms of :etardance; no mixing present in this representation.

If the exponential terms are to equal the phase and retardance directly, (not be functions of the phase and retardance.) they must be factored into the present form, such that all the amplitudes as a group multiply a single exponential which contains all the phase and retardance. An example of a function where the $\theta$ 's do not directly equal the phases and retardances is

$$
\rho_{0} e^{\theta_{0} \sigma_{0}}+\rho_{1} e^{\theta_{1} \sigma_{1}}+\rho_{2} e^{\theta_{2} \sigma_{2}}+\rho_{3} \mathrm{e}^{\theta_{3} \sigma_{3}}
$$

Since there is a product occuring between the amplitude terms and exponential phase terms in the exponential phase polarization aberration expansion, there are order dependent terms present. The product of the $\rho_{1} \sigma_{1}$ and $\theta_{2} \sigma_{2}$ terms yield

$$
\begin{aligned}
\left(\sigma_{0} \rho_{0}+\sigma_{1} \rho_{1}\right) \exp \left(j \theta_{2} \sigma_{2}\right) & =\left(\sigma_{0} \rho_{0}+\sigma_{1} \rho_{2}\right)\left(\sigma_{0} \cos \theta_{2}+j \sigma_{2} \sin \theta_{2}\right) \\
& =\sigma_{0} \rho_{0} \cos \theta_{2}+\rho_{1} \sigma_{1} \cos \theta_{2}+j \rho_{0} \sigma_{2} \sin \theta_{2}-\rho_{1} \sigma_{3} \sin \theta_{2}
\end{aligned}
$$

The first three terms contain the appropriate normalization, linear polarization and diagonal retardance contributions. The last term, a circular polarization term (of higher order assuming $\rho_{1}$ and $\theta_{2}$ small,) is a result of the multiplication of $\sigma_{1} \sigma_{2}$, and does not relate to the phenomena described by $\rho_{1}$ and $\boldsymbol{\theta}_{2}$. This is an order dependent term that results from multiplying Pauli spin matrices.

The order dependent terms are eliminated by symmetrizing the product. The next equation, the symmetrized exponential phase polarization aberration expansion, removes the order dependent terms by multiplying amplitudes and phases in both permutations,

$$
\begin{aligned}
\mathrm{J}(\mathrm{H}, \rho, \phi)= & 1 / 2\left[\left(\rho_{0} \sigma_{0}+\rho_{1} \sigma_{1}+\rho_{2} \sigma_{2}+\rho_{3} \sigma_{3}\right) \mathrm{e}^{\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1} \stackrel{\left.\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)}{ }\right.}\right. \\
& \left.+\mathrm{e}^{\left(\theta_{0} \sigma_{0}+\theta_{1} \sigma_{1}+\theta_{2} \sigma_{2}+\theta_{3} \sigma_{3}\right)}\left(\rho_{0} \sigma_{0}+\rho_{1} \sigma_{1}+\rho_{2} \sigma_{2}+\rho_{3} \sigma_{3}\right)\right] .
\end{aligned}
$$

No coupling occurs between noncommuting basis states. Each coefficient describes
only the amplitude or phase characteristics of a single basis state. For example, $\rho_{1}$ and $\theta_{2}$ do not generate a $\sigma_{3}$ component. The expansion is clean.

## The E Vector

Placing the phases and retardance in the exponential proved to be useful. An algebra will be explored which places all terms in the exponent. Define the " E vector" $\vec{E}$ corresponding to the $C$ vector $\vec{C}$ as

$$
\vec{C}=c_{0} \sigma_{0}+c_{1} \sigma_{1}+c_{2} \sigma_{2}+c_{3} \sigma_{3}=e^{e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3}}=e^{\vec{E}} .
$$

The E vector is the "logarithm" of the C vector and the C vector is the "exponential" of the E vector.

## Order Independence

Expanding the E vector, the following terms are obtained.

$$
\begin{aligned}
J(H, \rho, \phi)= & \sigma_{0}+\left(e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3}\right)+\left(e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3}\right)\left(e_{0} \sigma_{0}+e_{1} \sigma_{2}+e_{2} \sigma_{2}+e_{3} \sigma_{3}\right) \\
& +\left(e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3}\right)\left(e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3} y e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3}\right)+\ldots .
\end{aligned}
$$

All order dependent terms cancel as was found with the exponential phase expansion. For every order dependent term

$$
\mathrm{e}_{\alpha} \sigma_{\alpha} \mathrm{e}_{\beta} \sigma_{\beta}=j \mathrm{j}_{\alpha} \mathrm{e}_{\beta} \sigma_{\gamma} .
$$

there is a second term

$$
\mathrm{e}_{\beta} \sigma_{\beta} \mathrm{e}_{\alpha} \sigma_{\alpha}=-\mathrm{j}_{\alpha} \mathrm{e}_{\beta} \sigma_{\gamma}
$$

which cancels it.

## Relation to the C Vector

For small E vectors such that

$$
e_{0}, e_{1}, e_{2}, e_{3} \ll 1,
$$

then,

$$
c_{1} \cong e_{1}, c_{2} \cong e_{2}, c_{3} \cong e_{3} \quad \text {, and, } \quad c_{0}-1 \cong e_{0}
$$

This follows from the first order Taylor series expansion for the $E$ vector given above. This relationship is the matrix equivalent to the relation

$$
\ln (1+z) \cong z .
$$

The exact relationship between the C and E vectors is derived as follows. Consider the matrix

$$
e=e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3} .
$$

Find the similarity transform $S$ which diagonalizes $e$ and places its eigenvalues $\lambda_{1}$ and $\lambda_{2}$ on the diagonal,

$$
\Lambda=\left[\begin{array}{ll}
\lambda_{1} & 0 \\
0 & \lambda_{2}
\end{array}\right]=S^{-1} e S
$$

where

$$
\lambda_{1}, \lambda_{2}=e_{0} \pm \sqrt{e_{1}^{2}+e_{2}^{2}+e_{3}^{2}} .
$$

So

$$
e=S \Lambda S^{-1}
$$

Expanding the e vector

$$
\begin{aligned}
e^{e} & =e^{S \Lambda S^{-1}} \\
& =\sigma_{0}+S \Lambda S^{-1}+\frac{\left[S \Lambda S^{-1}\right)^{2}}{2!}+\ldots \\
& =\sum_{n=0}^{\infty} \frac{\left[S \Lambda S^{-1}\right]^{n}}{n!} \\
& =S\left[\sum_{n=0}^{\infty} \frac{\Lambda^{n}}{n!}\right] S^{-1}
\end{aligned}
$$

$$
\begin{aligned}
& =s e^{\Lambda} s^{-1} \\
& =s\left[\begin{array}{cc}
e^{\lambda_{1}} & 0 \\
0 & e^{\lambda_{2}}
\end{array}\right] \mathbf{s}^{-1} .
\end{aligned}
$$

Through this final expression, the e vector can be equated to conventional matrix expressions or C vectors.

Likewise let $\eta_{1}$ and $\eta_{2}$ be the eigenvalues of the matrix $c$,

$$
\begin{aligned}
\mathbf{c} & =c_{0} \sigma_{0}+c_{1} \sigma_{1}+c_{2} \sigma_{2}+c_{3} \sigma_{3} \\
\xi_{1}, \xi_{2} & =c_{0} \pm \sqrt{c_{1}{ }^{2}+c_{2}{ }^{2}+c_{3}{ }^{2}}
\end{aligned}
$$

aind the similarity transform $\mathbf{T}$ which diagonalizes $\mathbf{c}$,

$$
Z=\left[\begin{array}{ll}
\xi_{1} & 0 \\
0 & \xi_{2}
\end{array}\right]=\mathbf{T c} \mathbf{T}^{-1} .
$$

The E vector corresponding to the C vector is given by the transformation

$$
\begin{aligned}
c & =e^{e}=\exp \left[T \ln \Xi T^{-1}\right] \\
& =\exp \left[T\left[\begin{array}{cc}
\ln \xi_{1} & 0 \\
0 & \ln \xi_{2}
\end{array}\right] T^{-1}\right] .
\end{aligned}
$$

So

$$
\mathbf{e}=\mathbf{T}\left[\begin{array}{cc}
\ln \xi_{1} & 0 \\
0 & \ln \xi_{2}
\end{array}\right] \mathbf{T}^{-1} .
$$

## Nondiagonalizability

Two by two matrices which have two equal eigenvalues cannot be diagonalized by a similarity transform and are said to be nilpotent. Thus $C$ vectors and $E$ vectors for which

$$
c_{1}^{2}+c_{2}^{2}+c_{3}^{2}=0, \text { or, } e_{1}^{2}+e_{2}^{2}+e_{3}^{2}=0
$$

are not diagonalizable and cannot be converted between the $C$ vector and $E$ vector forms by the similarity transform method of the last section.

## Multiplication of $\mathbf{E}$ vectors

The following identity regarding the multiplication of $E$ matrices has been derived

$$
\mathbf{e}^{\vec{E}_{1}} \vec{e}^{2}=e^{\vec{E}_{1}+\vec{E}_{2}+\left[\vec{E}_{r} \overrightarrow{E_{2}}\right] / 2+\left[\vec{E}_{r}-\vec{E}_{2}\left[\vec{E}_{1} \vec{E}_{2}\right]\right] / 12+\ldots}
$$

where [ $\mathrm{a}, \mathrm{b}$ ] is the commutator

$$
[a, b] * a b-b a .
$$

The first two terms

$$
e^{\mathbf{e}_{1}}+\overrightarrow{E_{2}}
$$

are the order independent product of $\vec{E}_{1}$ and $\vec{E}_{2}$. The next term

$$
\left.\mathrm{e}^{\left[\overrightarrow{\mathrm{E}}_{\mathrm{r}}\right.} \overrightarrow{\mathrm{E}}_{2}\right] / 2
$$

is the first order dependent correction followed by order dependent corrections at still higher order. This expression converges rapidly for small $\vec{E}_{1}$ and $\vec{E}_{2}$.

## Divergence

Just as $\ln (x)$ diverges for $x \rightarrow 0$, the $E$ vector diverges as $J \rightarrow 0$. This is not a problem for transparent systems but presents mathematical complications for expansions about J $\cong 0$.

## The Exponential Polarization Aberration Expansion

The "exponential polarization aberration expansion" describes the instrumental polarization function in $E$ vector format. This expansinn takes the form

$$
J(H, \rho, \phi)=\exp \left[\sum_{k=0}^{3} e_{k}(H, \rho, \phi) \sigma_{k}\right]=e^{e_{0} \sigma_{0}+e_{1} \sigma_{1}+e_{2} \sigma_{2}+e_{3} \sigma_{3}} .
$$

The exponential polarization aberration expansion is order independent without requiring symmetrization between the amplitude and phase terms, as the exponential phase polarization aberration expansion required. It is convienient to keep the real (amplitude) and imaginary (phase) parts together in the exponent. This equation does have two disadvantages, divergence and nondiagonalizability of nilpotent matrices. It also has complications at higher order regarding the placement of amplitude and phase in the exponential.

## CHAPTER 9

## FINAL COMMENTS


#### Abstract

Summary This dissertation contains a method for treating the polarization effects of coatings or other weak polarizers in a manner analogous to the methods of aberration theory. This method calculates the instrumental polarization function $\mathbf{J}(\mathrm{H}, \rho, \phi)$ for radially symmetric optical systems with homogeneous and isotropic coatings on the interfaces. It is shown that the Taylor series expansion of the amplitude transmission functions of coatings is directly related to the amount of the second order polarization aberration of the system.

The limitations of the present methods are that it is limited to weak polarizers, that it only handles radially symmetric optical systems, and that the theory only handles coating effects to second order. Further, instrumental polarization arises from both interfaces and from propagation, but only interface effects are treated in this work. This treatment is also strictly monochromatic. The instrumental polarization arising from coatings is strongly wavelength dependent so a polychromatic formulation is desirable extension.

The highly transparent radially symmetric systems treated here are "best case" systems with small amounts of instrumental polarization. In these systems the instrumental polarization is small because the rays are near normal incidence and because the most common thin films, antireflection coatings or reflection enhancing coatings, are fairly low polarization coatings near their design wavelength.


Examples of systems not treated here are those containing gratings, holograms, electro-optic crystals. fold mirrors and other elements with large exgles of incidence. These are optical elements which are difficult to incorporate into geometrical aberration theory, because they are off-axis or because the polarization elements require higher order terms for useful characterization.

The method of polarization aberrations suggests a proceedure for selecting coatings to control and reduce the residual polarization of the optical system. First, coatings with small instrumental polarization should be selected; in particular $d_{12}$ should be small. Then, carefully chosen coatings with opposite signs of $d_{12}$ placed on different surfaces can compensate for coating induced instrumentai poiarization among the interfaces.

## Strong Polarizers

Strong polarizers are polarizers and retarders which substantially change the polarization state of light. Examples of strong polarizers are: 1) the "standard" linear polarizer with principal transmittances $\mathbf{k}_{1} \cong 1$ and $\mathbf{k}_{2} \cong 0$, and, 2) retarders with retardances greater than one radian, such as quarter wave and half wave retarders. The present polarization aberration theory is readily extensible to systems containing strong polarizers at normal incidence to the optical axis given a suitable Taylor series expansion of the Jones matrix of the strong polarizer or retarder. For transparent systems, the Jones matrix down the axis is a constant times an identity matrix, and the aberration expansion is performed about that matrix. For strongly polarizing systems, the Jones matrix can be expanded about the Jones matrix down the axis for the system in the same fashion, only the expansion is no longer about the identity matrix. For a polariscope with crossed polarizers, the expansion is about the zero matrix.

Strong polarizers have large order depeñueñt terms and care must be taken to use only meaningful approximations.

One potentially important application of polarization aberration theory is to the propagation of spherical waves through electro-optic media and modulators. Many modulators, such as the Kerr cell and Pockels cell, depend on polarization for their operation. These devices are usually operated with light beams which are nearly collimated. For a finite object size, there must be collimated beams over a spread of angles of incidence through the device. These beams experience a range of polarization response from the modulator and the polarization state varies across the image. Beyond a critical angle of incidence, the modulator may no longer satisfy the polarization requirements for the application.

Further, it is not always practical to use devices only in collimated light. This may require additional optical elements or extra space. With an expensive crystal in a production optical system, it would be desirable to squeeze as much light through a small crystal by focusing through it. A polarization aberration treatment of modulator performance would aid in design decisions assessing the effect iuiaving large ranges of angle of incidence and focusing through devices.

## Order Dependence

The single most difficult and time consuming aspect of this work has been handling the order dependence of polarizats. Multiplying matrices is straightforward but understanding the associated noncommutative algebra is not. The order dependence becomes a sizeable complication in continuing the aberration expansion to higher order.

## Names of Polarization Aberrations

The present naming convention for the polarization aberration terms is not eatirely satisfactory. These names are used here to emphasize the connection between the polarization aberrations and the geometrical aberrations. This naming convention confuses matters because the "defocus terms" other than wavefront defocus (such as linear polarization defocus) do not relate to defocus; no quadratic variation of phase is involved. All the defocus and the linear polarization defocus terms share is a $p^{2}$ aperture dependence.

If this form of polarization aberration analysis becomes widespread, improved nomenclature is required. I would like to suggest the following names:

| Quadpol | Linear Polarization Defocus | $\rho^{2}\left(\sigma_{1} \cos 2 \phi+\sigma_{2} \sin 2 \phi\right)$ |
| :--- | :--- | :--- |
| Linpol | Linear Polarization Tilt | $\mathrm{H} \rho\left(\sigma_{1} \cos \phi+\sigma_{2} \sin \phi\right)$ |
| Conpol | Linear Polarization Constant Piston | $\mathrm{H}^{2} \sigma_{1}$ |
| Quarticpol | Linear Polarization Spherical Aberration | $\rho^{4}\left(\sigma_{1} \cos 2 \phi+\sigma_{2} \sin 2 \phi\right)$ |
| Quadtard | Linear Retardance Defocus | ${\mathrm{j} \rho^{2}\left(\sigma_{1} \cos 2 \phi+\sigma_{2} \sin 2 \phi\right)}_{\text {Liniard }}$ |
| Contard | Linear Retardance Tilt | $\mathrm{jH} \rho\left(\sigma_{1} \cos \phi+\sigma_{2} \sin \phi\right)$ |
| Quartictard | Linear Retardance Constant Piston | $\mathrm{jH}^{2} \sigma_{1}$ |
|  | Linear Retardance Spherical Aberration | $\mathrm{j} \rho^{4}\left(\sigma_{1} \cos 2 \phi+\sigma_{2} \sin 2 \phi\right)$ |

## Polarization Ray Tracing

Polarization ray tracing is an alternative method to polarization aberration theory for calculating the instrumental polarization function $J(\vec{h}, \vec{p}, \lambda)$. Polarization ray tracing supplements the equations of ray tracing with polarization calculations. Polarization ray tracing is the direst attack on the calculation. no expansions, no approximations. Just select a ray path through the opticai system, calculate the
polarization matrix associates with each interface and each ray segment, and multiply the matrices together.

Consider a ray with object coordinate $\overrightarrow{\mathrm{h}}_{0}$ and pupil coordinate $\vec{\rho}_{0}$. Let Q be the number of optical interfaces in the system. $i_{q}$ be the angles of incidence of the ray. and $\theta_{\mathrm{q}}$ be the orientations of the plane of incidence fcr the ray; $\mathrm{i}_{\mathrm{q}}$ and $\theta_{\mathrm{q}}$ are obtained from a conventional ray trace. The Jones matrix asseciated with the ray at interface $\mathbf{q}$ is $\mathbf{J}_{\mathbf{q}}\left(i_{q}, \theta_{\mathbf{q}}\right)$. The Jones matrix for propagation from interface $\mathbf{q}$ to $\mathbf{q}+1$ is $\mathbf{J}_{\mathrm{q}+1, \mathrm{q}^{-}}$. The Jones matrix for the ray from object space $(\mathrm{q}=1$ ) to image space ( $\mathrm{q}=\mathrm{Q}$ ) is

$$
J\left[\vec{h}_{0, p_{0}}\right]=J_{Q} J_{Q, Q-1} J_{Q-1} J_{Q-1, Q-2} \ldots j_{2} j_{2,1} \mathbf{J}_{1}=\prod_{q=Q,-1}^{1} J_{q} J_{q, q-1}
$$

This polarization ray trace proceedure samples the value of the instrumental polarization function, $J\left[\vec{h}_{0}, \vec{\rho}_{0}, \lambda_{0}\right]$ at one point, $\left[\vec{h}_{0}, \vec{P}_{0}, \lambda_{0}\right]$. This proceedure is repeated for as many $\overrightarrow{\mathrm{h}}$ 's. $\vec{\rho}$ 's and $\lambda$ 's as are needed. By this méthod, ray tracing calculations and instrumental polarization calculations can be performed simultaneously if routines are included in the optical design program to calculate the interface polarization matrices $\mathbf{J}_{\mathbf{q}}$, and the propagation polarization matrices $\mathbf{J}_{\mathbf{q}, \mathbf{q}-1}$.

## Comparison of Polarization Aberrations and Polarization Ray Tracing

This research started as an effort to write a polarization ray tracing program to evaluate the effects of thin film coatings on the propagation of polarized light. It soon became apparent what an ambitious undertaking it was to merge thin film and ray tracing calculations, and particularly how little understanding there was of the form of the spatial variations of polarization induced by interfaces. Rather than continue to program software which calculated something ! didn't understand (the
spatial variation of instrumental polarization) I changed techniques. I developed the polarization aberration method to understand the basic forms of the spatial variation of instrumental polarization.

The polarization ray tracing program was excellent at producing numbers, large files full of them. Grids of rays were traced through optical systems and Mueller polarization matrices were calculated along each ray. (Meuller matrices wert being used because we wanted to be sure that we could handle incoherent as well as coherent light. We the contract sponsors, my research associates and myself) didn't realize at the time that the since none of our interface models displayed depolarization, that the much simpler Jones calculus was sufficient to handle the propagation of incoherent as we!! as coherent light through non-scattering polarizers. Thus, life and programming were more difficult than necessary due to the use of the Mueller calculus. For a grid of $\mathrm{N} \times \mathrm{N}$ rays, $16 \mathrm{~N}^{2}$ numbers were necessary to characterize the instrumental polarization function for one object and wavelength, and 16 FL N numbers to characterize F objects and L wavelengths. The polarization ray tracing method required large amounts of additionai programming to make any sense of all these numbers, particularly since I wasn't sure at the time what to expect.

The polarization aberrations occured as a means of understanding the undeilying optics behind the spatial variation of instrumental polarization and its connection to the angle of incidence function and the angle of incidence dependence of coatings. Although the polarization aberrations involve a large number of coefficients, it is orders of magnitude less numbers that are involved in a polarization ray tracing evaluation. The polarization aberrations yield direct understanding of the interaction of the coating with spherical waves at an interface.

Polarization ray tracing provides a more accurate calculation, and with sufficient programming, could give just as complete a picture of the insirumental polarization. Furthermore, a polarization ray tracing program can handle arbitrary systems for which it would be difficult to obtain reliable expansions, such as diffraction gratings on potato chip surfaces or aspheric lenses formed from birefringent crystals.

I would summarize this comparison by paraphrasing a statement that has often been made about classical aberration theory and ray tracing. Polarization aberrations are more suited for those who need to understand instrumental polarization and to develop creative stratages to reduce their effect. Polarization ray tracing is more suited for accurate analysis and for brute force attacks on instrumental polarization. Both methods have their place, but polarization ray tracing will probably overshadow polarization aberrations after enough insight has been giemed from polarization aberrations about the nature of the real problem and after the big polarization ray tracing software programs have been written and mastered.

Perhaps it will soon be possible, even easy, to readily assess the effect of a set of thin film coatings on the propagation of polarized light through optical systems.

## APPENDIX A

## NOTATION

The following notation conventions are adhered to throughout this work. Symbols which occur briefly are not included.

| $\mathrm{A}_{\mathrm{k}, \mathrm{u}, \mathrm{v}, \mathrm{w}}$ | Amplitude polarization aberration coefficient, |
| :---: | :---: |
| ${ }^{\text {app }}$ | Complex amplitude transmittance, p component, |
| $\mathrm{a}_{\text {S }}$ | Complex amplitude transmittance, s component, |
| $\mathrm{a}_{\mathrm{p}}$ | Modulus of the amplitude transmittance, p component, |
| $\mathrm{a}_{\text {S }}$ | Modulus of the amplitude transmittance, s component, |
| $\overline{\mathrm{a}}$ | Average amplitude transmission modulus, |
| B | Subscripted, basis Jones matrix, |
| $\vec{C}$ | C vector, |
| c | Subscripted, C vector elements, |
| c | Curvature of interface, |
| d | Normalized C vector elements in s and p coordinates, |
| d | Thickness of a single layer thin film, |
| $\overrightarrow{\mathrm{E}}(\mathrm{t})$ | Electric field, |
| $\overrightarrow{\mathrm{E}}$ | E vector, |
| e | Subscripted, E vector element, |
| G | Normalized x object height, |
| $\hat{H}$ | Jones vector for horizontal linearly polarized light, |
| H | Normalized y object height, |


| $\overrightarrow{\mathrm{h}}$ | Object vector, |
| :---: | :---: |
| I | Subscripted, intensity transmission coefficient, |
| i | Angle of incidence, |
| ${ }^{\prime}$ | Angle of refraction, |
| J | Jones matrix. |
| J | Jones vector, |
| $\mathbf{J}(\overrightarrow{\mathrm{h}}, \overrightarrow{\mathrm{p}})$ | Instrumental polarization function, |
| $\mathrm{J}_{\text {q }}$ | Jones matrix for q'th interface, |
| j | The imaginary number, $\sqrt{ } \mathbf{- 1}$. |
| j | Double subscript. Jones matrix element, |
| K | Beam crossection ratio for refraction. |
| x | Rasis coupling matrix, |
| k | Absorption coefficient, imaginary part of refractive index, n - jk , |
| $\hat{\mathbf{k}}$ | Normalized wavevector of light, |
| L | Optical path length from object to image, |
| $\hat{\mathbf{L}}$ | Jones vector for left circularly polarized light. |
| L | Jones matrix associated with a ray segment between surfaces. |
| $\overrightarrow{\mathrm{L}}$ | $C$ vector for a linear polarizer, |
| 1 | Left circular polarized transmission coefficient, |
| $\ell$ | Length of ray path segment. |
| M | Mueller matrix, |
| $\hat{\mathrm{m}}$ | Unit vector in the plane of incidence perpendicular to $\hat{\mathrm{n}}$. |
| N | Relative refractive index, $\frac{\mathrm{n}}{\mathrm{n}^{\prime \prime}}$ |
| n | Refractive index, |


| n | Slope of surface normal, |
| :---: | :---: |
| $\mathrm{n}^{\prime}$ | Refractive index following surface, |
| $\hat{\mathbf{n}}$ | Unit normal to surface, |
| $\mathrm{O}\left\{\mathrm{i}^{\mathrm{n}}\right\}$ | Terms of order $\mathrm{i}^{\mathbf{n}}$ and higher, |
| $\mathrm{O}\{\mathrm{f}(\mathrm{i}), \mathrm{n})$ | Function, returns terms of $f(i)$ of order less than or eaual to $n$, |
| $\mathrm{P}_{\mathrm{k}, \mathrm{u}, \mathrm{v}, \mathrm{w}}$ | Complex polarization aberration coefficient, |
| $\hat{\mathrm{p}}$ | Unit vector in p plane, |
| Q | Exponent, H•H power dependence, |
| Q | Subscript, Total number of interfaces |
| q | Subscript, interface numbering index, |
| R | Jones vector for right circularly polarized light, |
| R | Subscripted, power reflection coefficient, |
| R | Exponent, $\rho \cdot \rho$ power dependence, |
| R | Rotation matrix for polarization matricies, |
| R | Rotation matrix for C vectors, |
| $\overrightarrow{\mathrm{R}}$ | C vector for a retarder, |
| r | Right circular polarized transmission coefficient, |
| r | Subscripted, interface amplitude reflection coefficient, |
| S | Exponent, H- $\mathrm{\rho}$ power dependence, |
| S | Similarity transform matrix, |
| s | Subscript, component perpendicular to the plane of incidence, |
| $s$ | Subscript, exponent, $\cos \phi$ power dependence of an aberration term, |
| s | Slope of lines in y y-bar diagram, |
| $\hat{\mathbf{s}}$ | Unit vector in the s direction, tangential to surface, |
| T | Subscripted, power transmission coefficient, |
| t | Subscripted, interface amplitude transmission coefficient, |


| t | Time |
| :---: | :---: |
| u | Paraxial ray angle. |
| u | Exponent, subscript. H dependence of an aberration term: |
| $\hat{\mathbf{V}}$ | Jones vector for vertical linearly polarized light, |
| v | Exponent, subscript. $\rho$ dependence of an abersation term. |
| $\mathrm{W}(\overrightarrow{\mathrm{h}}, \vec{\rho})$ | Wavefront aberration function, |
| W | Phase shift measured in waves, |
| $\mathrm{w}_{\mathrm{u}, \mathrm{v}, \mathrm{w}}$ | Wavefront aberration coefficieni, |
| w | Exponent, subscript, $\phi$ dependence of an aberration term, |
| $\hat{\mathrm{x}}$ | $x$ unit vector. |
| x | Cartesian pupil coordinate, normalized. |
| x | Subscripted, paraxial ray x height. |
| $\hat{\mathbf{y}}$ | $y$ unit vector. |
| y | Cartesian pupil coordinate, normalized. |
| y | Subscripted, paraxial ray y height, |
| $\alpha$ | Direction cosine. |
| $\beta$ | Direction cosine, |
| $\beta$ | Film phase thickness, |
| $\Gamma$ | Lagrange invariant, |
| $\gamma$ | Taylor series order, |
| $\Delta$ | Mean phase change, |
| $\delta$ | Phase, |
| $\delta$ | $\delta_{S}-\delta_{p}$. birefringence. |
| $\delta_{\text {c }}$ | Circular birefringence. |


| $\delta_{1}$ | Linear birefringence, |
| :---: | :---: |
| $\delta_{p}$ | $S$ component phase change across interface, |
| $\delta_{s}$ | P component phase change across interface, |
| $\delta_{0}$ | Average phase delay, |
| $\epsilon$ | A phase, |
| $\epsilon(t)$ | Time dependent phase, |
| $\theta$ | Orientation of plane of incidence, measured in radians, ccunterclockwise from y-axis. |
| $\theta$ | Arctan of ratio of polarization ellipse axes, |
| $\theta$ | Subscripted, exponential phase coefficient. |
| $\Lambda$ | Eigenvalue matrix, diagoinal, |
| $\lambda$ | Wavelength, |
| $\lambda$ | Subscripted, eigenvalues, |
| $\rho$ | Radial pupil coordinate, normalized, |
| $\vec{p}$ | Pupil vector to ray coordinate in pupil, |
| $\rho$ | Amplitude part of a complex number, |
| $\sigma_{0}$ | İdentity mairix, $2 \times 2$. |
| $\sigma_{1}, \sigma_{2}, \sigma_{3}$ | Pauli spin matrices, |
| $\Upsilon$ | Net optical amplitude transmittance along the axis through the system, |
| $\tau$ | Amplitude transmittance at normal incidence. $7=a_{S}(0)=a_{p}(0)=c_{0}$ |
| $\tau$ | Linear dichroism, |
| $\tau$ | Amplitude transmittance for a ray path segment, |
| $\Phi_{\mathrm{k}, \mathrm{u}, \mathrm{v}, \mathrm{w}}$ | Phase polarization aberration coefficient, |
| $\phi$ | Angular pupil coordinate, measured in radians, counterciockwise from y -axis, |

$\phi \quad$ Phase part of a complex number,
$\psi \quad$ Angular orientation of major axis of polarization ellipse.
Mean optical frequency,
Jones vector for $\mathbf{+ 4 5}$ degree linearly polarized light.
Jones vector for -45 degree linearly polarized light.

## Subscripts

| c | Chief Ray, |
| :--- | :--- |
| e | Stop, |
| k | Polarization type: 0,1,2,3, |
| m | Marginal ray, |
| p | P plane, plane of incidence, |
| q | Surface index, |
| s | S plane, perpencidular to the plane of incidence,, |
| u | H dependence, |
| v | $\rho$ dependence, |
| w | $\phi$ dependence,, |
| x | x comporent, |
| y | y component. |
| $\gamma$ | Taylor series order,, |

## Conventions

## Boldface Matricies

Primes, '. refer to quantities after an interface.
[ ] Square brackets, vector or matrix,
$\rightarrow \quad$ Vectors,
^ Unit vectors.
|| $\vec{x}|\mid \quad$ Norm of $\vec{x}$.

## APPENDIX B

## A COMPARISON OF THE JONES AND MUELLER CALCULUS

This appendix contains a comparison of the Jones and Mueller calculus and discusses the role of depolarization in optical design.

There are two principal computational methods for treating polarization problems, the Jones calculus and the Mueller calculus. Both calculi were developed in Cambridge. Massachusetts in the 1940's. This dissertation relies exclusively on the Jones calculus for developing the polarization aberrations because I have found the problem much easier to iormulate with the Jones calculus. However, the Jones calculus will not readily treat problems involving the depolarization of light while the Mueller calculus will, so I have included this comparison of the calculi. Using the equations contained herein, all Jones matrix results can be converted to the Mueller formalism.

## The Mueller Calculus

ln the Mueller calculus the state of polarization of light is described by the Stokes vector, a four element real vector, Stokes (1852), Shurcliff (1961). Polarizers are characterized by a four by four element real matrix, the Mueller matrix, Mueller (1946. unpublished), Mueller (1948), Parke (1949). The Mueller calculus will not be described in detail here. Introductions to the Mueller calculus are found in Shurcliff (1961, 8.2), Theocaris and Gdoutos (1979, 4.3.4), Azzam and Bashara (1977. 2.12), and Girrard and Burch (1975, IV.3).

## Depolarization

Depolarization can be operationally defined as any optical system effects which cause some fraction of a completely polarized incident beam to become partially polarized. Examples of depolarizing effects are: scattering from very rough surfaces, reflecting from surfaces covered with scratches and dirt, transmission through a turbid (scattering) medium, and transmission through a medium with a rapidly varying polarization or retardance. The depolarizing tendency of most high quality polarizers is so small as to be negligible in most applications, Shurcliff (1961 pg.33).

The frequent use of the term "depolarization" to describe any change in the state of polarized light is incorrect, unless it specifically involves the coupling of completely polarized light into unpolarized light. For example, the change in polarization of a linearly polarized coherent light beam on reflection from a metallic surface at non-normal incidence is not depolarization. The reflected beam is elliptically polarized and, in the absence of scattering, it remains completely polarized. The correct terms for such changes of polarization state are "polarization coupling" and "polarizatioñ rotation".

Optical fibers typically display large amounts of depolarization and polarization coupling, particularly at bends in the fiber.

## What the Jones and Mueller Calculus Describe

A quasimonochromatic optical field can be described by five quantities when the higher order statistical properties of the light are ignored. A quasimonochromatic beam propagating in one direction can be decomposed into a completely polarized component and an unpolarized component. The completely polarized component has four parameters, which can be expressed by the amplitude
and phase of the $x$ and $y$ components of the electric field. These parameters comprise its Jones vector. The unpolarized component has one parameter, its intensity. The unpolarized component has no long term phase relationship or coherence with the completely polarized component. The Jones calculus treats only the four parameters of the polarized component. It does not keep track of the unpolarized component of the fields. The Mueller calculus treats the unpolarized component and three of the polarized components, the amplitudes in x and y and the difference in phase between the $x$ and $y$ components. It does not calculate the fourth polarized parameter, the absolute value of the phase. Thus the Mueller calculus will not, by itself, handle multiple coherent beams such as in interferometry. Likewise, the Jones calculus cannot, by itself, handle scattering and depolarization effects in optical systems.

## Measuring the Jones and Mueller Parameters

The Mueller matrix and Stokes vector for an optical element are measured using a radiometer and sets of polarizers and retarders. Gerrald and Burch, (1975, pg.202) and Theocaris and Gdoutos (1979, section 5.6 ) give procedures for measuring the elements of a Stokes vector and Mueller matrix. Likewise, the Jones matrix and Jones vector are determined by a similar set of polarization measurements. Gerrald and Burch give a procedure to determine the Jones matrix and Jones vector with a sequence of intensity measurements. However, from intensity measurements, the Jones matrix and Jones vector are only determined to within a factor of $\exp (\mathrm{j} \phi)$ where $\phi$ is the unknown absolute value of the phase of all the elements. This measurement is sufficient except for instruments such as the white light Michelson interferometer or a phased array optical system where the absolute phase, relative to a reference beam, is important. This absolute phase, $\phi$, is just the optical path length
through the system modulo the wavelength, which needs to be known to a small fraction of a wavelength to be useful. One technique which could measure this absolute phase is a Mach Zehnder interferometer wit? white light. The piece under test is placed in the test arm. The reference arm must then be precisely calibrated for iength so that is optical path length is accurately known.

## Mathematical Complexity

The Jones calculus is formulated in terms of the electric field amplitudes which cannot be measured directly. The Mueller calculus is a strictly empirical calcuius, depending only on observable quantities, i.e. intensity measurements. Thus, it is accurate to say that the Jones calculus is an "amplitude calculus" and the Mueller calculus is an "intensity calculus".

Mathematically, the Jones calculus is far simpler than the Mueller calculus. The Jones matrix has four complex elements, or eight degrees of freedom. Every Jones matrix corresponds to a unique and physically realizable polarization device. Since the Jones calculus can analyze the absolute phase, it will distinguish between polarizers of differing thickness which are otherwise identical. The Jones calculus will not characterize depolarizers.

The Mueller matrix with its sixteen real elements will describe depolarizers but will not distinguish between otherwise identical polarizers of differing optical path length. Furthermore, not every possible Mueller matrix corresponds to a physically realizable polarizer. Neither calculus is complete in describing pularizers and light at this level of statistical complexity of the optical fields. The Jones calculus is complete in describing the polarization properties of all possible nondepolarizing optical elements.

The Mueller calculus is an intensity calculus. Since the intensity is the amplitude squared, most mathematical expressions are far more complex expressed in the Mueller calculus than in the Jones calculus. In particular, trigonometric expressions become quite cumbersome when squared. The Jones calculus is far easier to manipulate, and thus provides insight more readily. For example, the Jones and Mueller matrices for a half wave retarder with the fast axis oriented at an angle $\theta$ are:

$$
\begin{gathered}
\mathbf{J}=\left[\begin{array}{cc}
-\cos 2 \theta & -\sin 2 \theta \\
-\sin 2 \theta & \cos 2 \theta
\end{array}\right], \text { and, } \\
\mathbf{M}=\left[\begin{array}{cccc}
1 & 0 & 0 & 0 \\
0 & \cos 4 \theta & \sin 4 \theta & 0 \\
0 & \sin 4 \theta & -\cos 4 \theta & 0 \\
0 & 0 & 0 & -1
\end{array}\right] .
\end{gathered}
$$

## Characterizing Depolarization

The Jones matrix is applicable only to polarizers which do not scatter or depolarize light. Thus it represents the idealization of polarizer behavior, as opposed to describing the typically undesirable feature of depolarization. It is sometimes said that the Jones calculus is for coherent light problems while the Mueller calculus is for incoherent problems. This is not quite correct but close. The Jones calculus is quite capable of handling problems involving incoherent or partially coherent light. see Jones (1947). What the Jones calculus cannot handle is polarizers which scatter or depolarize the light. If the light incident on a polarizer is completely polarized and the light transmitted by that polarizer is completely polarized, then that polarizer can be completely described by a Jones matrix (or a Mueller matrix). If some of the transmitted light is depolarized, then the Jones matrix can only characterize that portion of the light which remains completely polarized.

For simple ideal interfaces, including reflection, refraction and homogeneous thin film coatings, no depolarization is predicted by theory. No depolarization term is contained in the Fresnel equations, for example. A small depolarization occurs at these interfaces in practice because of surface roughness and possibly scratches and dirt. This depolarization component is kept small through careful optical fabrication and maintenance practices, and is not yet predicted in detail by theory. Indeed, depolarization generally gets worse with time as surfaces accumulate dirt or are abused. Thus the depolarization of an optical system can be considered as more of an empirical phenomenon than one accessible to a theoretical treatment in the optical design process. The depolarization component should be minimized by careful fabrication practices, and it is unusual that an analytical form would be known.

For a system where the presence of depolarization is important, the proper treatment would be to analyze the depolarization by measuring the Mueller matrices of the system and its component elements. It is not of great concern here that the Jones calculus does not treat depolarization. There is a second method, the Mueller calculus, which is principally empirical but also analytical, to handle this problem. The mathematical advantages of the Jones calculus, its mathematical simplicity in comparison with the Mueller calculus, and what is more important, its ability to handle the coherent addition of optical beams, far outweigh this particular shortcoming. Further, to formulate a theory of polarization aberrations in terms of the Mueller calculus involves unjustified and unnecessary mathematical complexity.

## Relationship Between the Jones Matrix and the Mueller Matrix

The mapping from the Jones matrix space to the Mueller matrix space is one to one or functional. The mapping from the Mueller matrix space to the Jones matrix space is such that only a subset of the Mueller matrices have exact
correspondence with a Jones matrix, and for this subset the mapping is many to one or relational.

These properties lead to a simple and direct set of equations to rewrite a Jones matrix as a completely equivalent Mueller matrix. As a practical matter, Mueller matrices can be converted to "nearest" Jones matrices if we first strip the depolarization component off the Mueller inatrix. Then, these "reduced" or nondepolarizing Mueller matrices can be converted into Jones matrices whose absolute phase is undetermined, but are otherwise unique.

Table 12 contains the equations to transform a Jones matrix into the equivalent Mueller matrix. This discussions of this transformation can be found in Gerrald and Burch (1975, equation F.5) or Thencaris and Gdoutos (1979, equation 4.4).

## Summary

In general. the Mueller calculus is to be preferred for experimental work and the Jones calculus for theoretical work. In experimental work the depolarization should be routineiy measured along with ail the other polarization parameters.

In the optical design of most instruments, the depolarization can be ignored, nct because it is unimportant, but because it is a separate issue from the wavefront aberrations and instrumental polarization. Depolarization, as a phenomenon, is much closer in spirit to stray light and scattering and is probably more deserving of an empirical approach, at least at this time, than a theoretical or optical design approach. This is not to say that depolarization effects should not be included in optical design along with the other instrumental polarization effects; I can make some suggestions for treating this problem. But depolarization is far afield from the mathematical treatment of instrumental polarization developed here. Throughout this

## TABLE 12

## TRANSFORMATION OF JONES MATRICES INTO MUELLER MATRICES

This set of equations transform the elements of a Jones matrix into the elements of the equivalent Mueller matrix.

$$
\begin{aligned}
& 2 M_{11}=J_{11}^{*} J_{11}+J_{21}^{*} \mathrm{~J}_{21}+\mathrm{J}_{12}^{*} \mathrm{~J}_{12}+\mathrm{J}_{22}^{*} \mathrm{~J}_{22} \quad . \\
& 2 M_{12}=J_{11}^{*} J_{11}+J_{21}^{*} J_{21}-J_{12}^{*} J_{12}-J_{22}^{*} J_{22} \text {. } \\
& 2 \mathrm{M}_{13}=\mathrm{J}_{11}^{*} \mathrm{~J}_{12}+\mathrm{J}_{21}^{*} \mathrm{~J}_{22}+\mathrm{J}_{12}^{*} \mathrm{~J}_{11}+\mathrm{J}_{22}^{*} \mathrm{~J}_{21} \text { 。 } \\
& 2 \mathrm{M}_{14}=\mathrm{j}\left(\mathrm{~J}_{11}^{*} \mathrm{~J}_{12}+\mathrm{J}_{21}^{*} \mathrm{~J}_{22}-\mathrm{J}_{12}^{*} \mathrm{~J}_{11}-\mathrm{J}_{22}^{*} \mathrm{~J}_{21}\right) \text {, } \\
& 2 M_{21}=J_{11}^{*} J_{11}+J_{12}^{*} J_{12}-J_{21}^{*} J_{21}-J_{22}^{*} J_{22} \text {, } \\
& 2 M_{22}=J_{11}^{*} J_{11}+J_{22}^{*} J_{22}-J_{21}^{*} J_{21}-J_{12}^{*} J_{12} \text {, } \\
& 2 M_{23}=J_{12}^{*} J_{11}+J_{11}^{*} J_{12}-J_{22}^{*} J_{21}-J_{21}^{*} J_{22} \text {, } \\
& 2 \mathrm{M}_{24}=\mathrm{j}\left(\mathrm{~J}_{11}^{*} \mathrm{~J}_{12}+\mathrm{J}_{22}^{*} \mathrm{~J}_{21}-\mathrm{J}_{21}^{*} \mathrm{~J}_{22}-\mathrm{J}_{12}^{*} \mathrm{~J}_{11}\right) \text {. } \\
& 2 M_{31}=J_{11}^{*} J_{21}+J_{21}^{*} J_{11}+J_{12}^{*} J_{22}+J_{22}^{*} J_{12} \text {. } \\
& 2 M_{32}=J_{11}^{*} J_{21}+J_{21}^{*} J_{11}-J_{12}^{*} J_{22}-J_{22}^{*} J_{12} \text {, } \\
& 2 M_{33}=J_{11}^{*} J_{22}+J_{21}^{*} J_{12}+J_{12}^{*} J_{21}+J_{22}^{*} J_{11} \text {. } \\
& 2 M_{34}=j\left(J_{11}^{*} J_{22}+J_{21}^{*} J_{12}-J_{12}^{*} J_{21}-J_{22}^{*} J_{11}\right) \text {. } \\
& 2 M_{41}=j\left(J_{21}^{*} J_{11}+J_{22}^{*} J_{12}-J_{11}^{*} J_{21}-J_{12}^{*} J_{22}\right) . \\
& 2 M_{42}=j\left(J_{21}^{*} J_{11}+J_{12}^{*} J_{22}-J_{11}^{*} J_{21}-J_{22}^{*} J_{12}\right) \text {. } \\
& 2 M_{43}=j\left(J_{21}^{*} J_{12}+J_{22}^{*} J_{11}-J_{11}^{*} J_{22}-J_{12}^{*} J_{21}\right) \text {. } \\
& \left.2 \mathrm{M}_{44}=\mathrm{J}_{22}^{*} \mathrm{~J}_{11}+\mathrm{J}_{11}^{*} \mathrm{~J}_{22}-\mathrm{J}_{12}^{*} \mathrm{~J}_{21}-\mathrm{J}_{21}^{*} \mathrm{~J}_{12}\right) \text {. }
\end{aligned}
$$

dissertation, all polarizers considered are non-depolarizing, or what might be termed, "coherent polarizers". Included here are all dichroic polarizers, retarders, reflections. refractions and thin film coatings for which scattering and depolarization can be safely neglected. With coherent. completely polarized light, the light remains coherent and completely polarized throughout the optical system. These optical system polarization effects can be completely characterized with the Jones matrix. and all results can be converted to Mueller matrices as needed.

## APPENDIX C

## THE MEANING OF THE ELEMENTS OF THE C VECTOR

This appendix discusses the representation of specific classes polarizers with $C$ vectors to elaborate on the meaning of the different elements of the $C$ vector. Table 9 lists the association between the forms of polarization behavior and the amplitude and phase of the elements of the $\mathbf{C}$ vector.

## Absorption

In simple absorption. both the x and y components of the amplitude are absorbed equally:

$$
a_{s}=a_{p}=\rho_{0} .
$$

The Jones matrix is a constant times the identity matrix,

$$
J=\rho_{0}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right] .
$$

This Jones matrix is expressed in exponential form as

$$
\mathrm{J}=\mathrm{e}^{\alpha l \sigma_{0}}
$$

where $\alpha$ is the absorption coefficient per unit length and $\ell$ is the path length. The $C$ vector for absorption is

$$
\vec{C}=\left[\rho_{0}, 0,0,0\right]=e^{\alpha l \sigma_{0}}[1,0,0,0]
$$

## Propagation, Phase Delay

For a simple propagation over length $\ell$. through free space or a medium with refractive index $n$, the optical path length $L$ is

$$
\mathbf{L}=\mathbf{n l} .
$$

The phase delay produced during propagation measured relative to the incident phase of the beam is

$$
\delta=\frac{2 \pi n l}{\lambda}
$$

For propagation, the amplitude transmission relations are:

$$
a_{p}=a_{S}=e^{j \delta}
$$

The Jones matrix is

$$
\mathbf{J}=\mathrm{e}^{\mathrm{j} \delta}\left[\begin{array}{ll}
1 & 0 \\
0 & 1
\end{array}\right]=\exp ^{\mathrm{j} \delta \sigma_{0}} .
$$

The $\mathbf{C}$ vector is

$$
\vec{C}=e^{j \delta}[1,0,0,0] .
$$

## Polarization

Polarization is the property of materials which decompose the incident light into two orthogonal polarization states and reflect, transmit, diffract or scatter these components with a different transmission coefficient. A technical distinction is made here between polarization and retardance, such that polarization is strictly the difference in transmission and retardance is the difference in phase. The word polarization is commonly used in less precise discussions to refer to both "polarization" and retardance effects.

Dichroism is the material property of being polarizing on transmission due to differential polarization state dependent absorption. Tourmaline, herapathite and
common sheet polarizer are examples of dichroic materials at normal incidence.
Refraction and reflection at nonnormal incidence display polarization due to the polarization differences in partial reflection. Since these polarization effects do not arise from differential polarization dependent absorption, they are not dichroic polarizers.

Polarization is classified as linear, circular or elliptical depending on the eigenpolarizations of the polarizer.

Linear polarization is mathematically represented as follows. Let the x and y axes be parallel to the eigenpolarizations. Let the amplitude transmittances be

$$
a_{x}=r_{1}, a_{y}=r_{2}
$$

where $r_{1}$ and $r_{2}$ are real. Then, the Jones matrix and $C$ vectors are,

$$
\left.\begin{array}{rl}
\mathbf{J} & =\left[\begin{array}{cc}
\mathbf{r}_{1} & 0 \\
0 & r_{2}
\end{array}\right], \\
\mathbf{C} & =\left[\rho_{0}, \rho_{1}, 0,0\right.
\end{array}\right] .
$$

The element $\rho_{0}=\left(r_{1}+r_{2}\right) / 2$ is the average amplitude transmittance; it is the square root of the transmittance of unpolarized light. The element $\rho_{1}=\left(r_{1}-r_{2}\right) / 2$ is the amount of linear polarization.

If the dichroic polarizer is rotated by $45^{\circ}$. the Jones matrix and $C$ vector become.

$$
\mathbf{J}^{\prime}=\mathbf{R}(\pi / 4) \mathbf{J} \mathbf{R}(-\pi / 4)=\frac{1}{2}\left[\begin{array}{ll}
r_{1}+r_{2} & r_{1}-r_{2} \\
r_{1}-r_{2} & r_{1}+r_{2}
\end{array}\right] \text {. }
$$

and

$$
\vec{C}=\left[\rho_{0}, 0, \rho_{2}, 0\right] .
$$

Again, $\rho_{0}=\left(r_{1}+r_{2}\right) / 2$ is the average transmittance without polarization. The amount of linear polarization is the same, $\rho_{2}=\left(r_{1}-r_{2}\right) / 2$, but it has been snifted to the $c_{2}$
element. The general form for the $\mathbf{C}$ vector of a linear partial polarizer with average transmittance $\rho_{0}$. and linear polarization $\chi$. oriented at an angle $\theta$ is

$$
\vec{C}=\left[\rho_{0}, \chi \cos 2 \theta, \chi \sin 2 \theta .0\right] .
$$

Circular polarization is the property of certain materials which decompose the incident light into left and right circular polarized components and transmit them with different transmission coefficients. If the amplitude transmission for right circularly polarized light is, $r$, and for left circularly polarized light is, 1 . the Jones matrix and $C$ vectors are,

$$
\mathbf{J}=\frac{r}{2}\left[\begin{array}{rr}
1 & j \\
-j & 1
\end{array}\right]+\frac{1}{2}\left[\begin{array}{rr}
1 & -j \\
j & 1
\end{array}\right]=\frac{1}{2}\left[\begin{array}{rr}
r+1 & -j(r-1) \\
j(r-1) & r+1
\end{array}\right] .
$$

and.

$$
\overrightarrow{\mathrm{C}}=\left[\rho_{0}, J, 0, \rho_{3}\right]
$$

Again, $\rho_{0}$ the $c_{0}$ element, is the average amplitude transmission, and $\rho_{3}=(r-1) / 2$, the $c_{3}$ element, is the amount of circular polarization. Due to the rotational invariance of $\sigma_{0}$ and $\sigma_{3}$, this matrix has the same form in rotated coordinates. This is because the decomposition of the incident light into circular polarized components in a polarizer is independent of the orientation of the polarizer.

Elliptical polarization is expressed as a combination of linear and circular polarization.

## Retardance

Retardance is the property of certain polarizing elements where the eigenpolarizations are transmitted with a relative phase shift. In birefringent materials, the two eigenpolarizations have different refractive indices and thus different optical path lengths through the material. Thus the slower eigenpolarization exits the element with a phase delay or retardation with respect to
the faster component. The birefringence of an element is usually specified as this phase delay expressed in fractions of a wavelength. Thus a quarter wave retarder has optical path lengths for the two eigenpolarizations which differ by one quarter of a wavelength of light. Interfaces between nonbirefringent materials cani also cause retardance at nonnormal incidence. Transmission through thin film coatings and reflection from metals are examples of such retardance or "induced birefringence".

Birefringent polarizing elements are usually called retarders. Retarders can be linear, circular or elliptical depending on the form of the eigenpolarizations. Quartz crystals will form all three types of retarder. A section of quartz cut parallel to the optic axis forms a linear retarder; a section cut perpendicular to the optic axis forms a circular retarder; oblique sections display elliptical retardance. Many liquids, such as turpentine or dextrose in water solution, as well as liquid crystals display circular birefringence. To correctly model and use a retarder, it is necessary to know which eigenpolarization leads and which is transmitted more slowly. For a linear retarder, the axis of the fast eigenpolarization is referred to as the "fast axis".

The amplitude transmittance relations for a linear retarder with fast and slow axes aligned with the x and y axes are:

$$
a_{x}=e^{-j \delta x}, \text { and, } a_{y}=e^{-j \delta y}
$$

Minus signs are used because all optical materials delay the phase of the transmitted light relative to a signal propagated in vacuum. The Jones matrix and C vectors for iinear retardance parallel to the $x$ and $y$ axes are

$$
J=\left[\begin{array}{ll}
e^{-j \delta_{x}} & 0 \\
0 & e^{-j \delta_{y}}
\end{array}\right]=e^{-j \delta_{0}} e^{-j \delta \sigma_{i}}
$$

and.

$$
\vec{C}=\mathrm{e}^{-\mathrm{j} \delta_{0}}\left[\cos \delta_{.}-\mathrm{j} \sin \delta_{1} 0.0\right] .
$$

Here the average phase is

$$
s_{0}=\left(\delta_{x}+\delta_{y}\right) / 2
$$

and the linear retardance is

$$
\delta=\left(\delta_{x}-\delta_{y}\right) / 2 .
$$

The $\mathbf{C}$ vector for linear retardance at $45^{\circ}$ to the axes is

$$
\vec{C}=e^{-\mathrm{j} \delta_{0}}[\cos \delta, 0,-\mathrm{j} \sin \delta, 0] .
$$

A circularly birefringent optical element with phase delays. $\delta_{\mathrm{r}}$ and $\delta_{1}$, has the Jones matrix,

$$
J=\frac{1}{2} e^{-j \delta_{r}}\left[\begin{array}{rr}
1 & -j \\
j & 1
\end{array}\right]+\frac{1}{2} e^{-j \delta_{I}}\left[\begin{array}{rr}
1 & j \\
-j & 1
\end{array}\right] .
$$

Since the average phase delay is

$$
\delta=\frac{\delta_{\mathrm{r}}+\delta_{1}}{2}
$$

and the circular retardance $\delta_{c}$ is

$$
\delta_{c}=\frac{\delta_{r}-\delta_{1}}{2} .
$$

the Jones matrix and C vector for a circular retarder are,

$$
J=e^{-j \delta}\left(\cos \delta_{c} \sigma_{0}+\sin \delta_{c} \sigma_{3}\right)=e^{-j \delta} e^{-j \delta_{c} \sigma_{3}} .
$$

and,

$$
む=e^{-j \delta}\left[\cos \delta_{c}, 0,0,-j \sin \delta_{c}\right] .
$$

Elliptical retardance is expressed as a combination of linear and circular retardance.

In summary, for all cases, the amplitude of the $\mathbf{C}$ vector element refers to amplitude effects: absorption, transmission or polarization. The phase of the C
vector elements refer to phase and retardance. Describiag a weak polarizer with the four complex coefficients of the $\mathbf{C}$ vector often proves far easier than manipulating Jones matrices.

## APPENDIX D

## PARAXIAL SREW RAYS

## Introduction

This appendix contains results from paraxial optics useful for the development of the polarization aberrations. First there is an explanation of the notation for the coordinate system and for paraxial rays. This is followed by the derivation of expressions for the angle of incidence and orientation of the plane of incidence for paraxial skew rays. Finally, equations are given for the marginal and chief ray angles of incidence as functions of paraxial ray trace and y y-bar diagram parameters.

## Cocrdinate System

The coordinate system adupted here is a normalized right handed coordinate system. The $z$ axis is the optical axis of a rotationally symmetric optical system. Light initially travels with a positive z component of the wave vector, i.e. it travels in the direction of increasing z . The y -axis is depicted as pointing upwards. Looking down the z -axis, clockwise rotation brings a line from the x -axis to the y axis.

Figure 10 shows the notation. G and H are the normalized object coordinates, G along the x -axis and H along the y -axis. Normalization is performed such that around the edge of a circular field of view,

$$
\sqrt{\mathrm{G}^{2}+\mathrm{H}^{2}}=1 .
$$

The object vector, $\vec{H}$ is defined as

$$
\overrightarrow{\mathrm{H}}=(\mathrm{G}, \mathrm{H}) .
$$

The pupil coordinates are $\rho$ and $\phi$ in polar form or $x$ and $y$ in Cartesian form. $\rho, \mathrm{x}$ and y are also normalized such that at the edge of a circular pupil

$$
\sqrt{x^{2}+y^{2}}=\rho=1 .
$$

$\phi$ is defined here as it is in much of geometric optics, and in defiance of most analytical geometry, as being zero on the y-axis and increasing counterclockwise. Then,

$$
x=\rho \sin \phi \text { and } y=\rho \cos \phi .
$$

The pupil vector is defined as

$$
\vec{\rho}=(\mathbf{x}, \mathbf{y}) .
$$

Objects are typically located on the y -axis such that $\mathrm{G}=\mathbf{0}$. Then the orientation of the plane of incidence, $\theta$, for the $y-z$ meridional plane is always vertical, $\theta=0$.

## Subscripts

The following subscript notation $s$ used for the various ray components. Subscripts c and m refer to quantities associated with the chief and marginal rays in the $y-z$ plane. Subscript $q$ refers to the quantity at the $q^{\text {th }}$ interface in the system. Where a set of expressions refers to a single surface, the surface subscript is omitted. Subscript e is the quantity evaluated at the system stop (mnemonic entrance pupil). For skew rays it is necessary to distinguish beiween quantities measured relative to the x -axis, subscript x , and y -axis, subscript y .

## Meridional Rays, y-z Plane

The slope of a ray, $u$, is positive if a counterclockwise rotation brings the axis to the ray. The slope of the surface normal, $n$, is likewise positive if clockwise
rotation brings the axis to the ray. Thus, for a paraxial spherical surface with curvature c .

$$
\mathrm{n}=-\mathrm{yc} .
$$

The definitions of $u$ and $n$ are consistent with the conventional definition of slope as

$$
\mathrm{m}=\frac{\mathrm{df}(\mathrm{y})}{\mathrm{dz}} .
$$

The angle of incidence, $i$, is defined as

$$
\mathbf{i}=\mathbf{u}-\mathbf{n} .
$$

i is positive if a counterclockwise rotation brings the surface normal to the ray.
The chief ray in the $y-z$ meridional plane is thie paraxial ray from the object point. $(G, H)=(0,1)$ through the center of the pupil, $(\rho, \phi)=(0,0)$. The height of the ray at the $q^{\text {th }}$ interface is denoted, $y_{c, q}$. Its angle of incidence measured from the normal is $i_{c, q}$. The orientation of the plane of incidence measured clockwise from the y axis is always $\theta_{\mathrm{c}, \mathrm{q}}=0$, since the ray is in the $\mathrm{y}-\mathrm{z}$ meridional plane.

The marginal ray in the $y-z$ meridional plane is the paraxial ray from the center of the object, $(\mathrm{G}, \mathrm{H})=(0,0)$, through the top of the pupil. $(\rho, \phi)=(1,0)$ or $(\mathrm{x}, \mathrm{v})=$ (0,1).

Meridional rays in the $\mathrm{x}-\mathrm{z}$ plane follow the same conventions with x substituted for $\mathbf{y}$.

## Paraxial Skew Rays

The paraxial ray trace follows from a linearization of Snell's Law. Because of the linearity of paraxial optics, any meridional paraxial ray can be expressed as the linear combination of two linearly independent paraxial rays in the same meridional plane. This linearity extends to paraxial skew rays but $x$ and $y$ components must be added separately. An arbitrary skew paraxial ray can be
expressed as the linear combination of any four linearly independent paraxial rays. The chief and marginal paraxial rays in the $x-z$ plane and the chief and marginal rays in the $y-z$ plane are used as the basis ray set. For a radially symmetric system, both the $x-z$ and $y-z$ chief and marginal ray parameters are determined from a paraxial ray trace calculation for a single plane because the x and y components of the ray intercept, ray slope and angle of incidence are equal:

$$
\begin{gathered}
x_{c, q}=y_{c, q}, x_{m, q}=y_{m, q} \\
u_{c, q}=u_{c, x, q}=u_{c, y, q}, u_{m, q}=u_{m, x, q}=u_{m, y, q} \\
i_{c, q}=i_{c, x, q}=i_{c, y, q}, \text { and } i_{m, q}=i_{m, x, q}=i_{m, y, q}
\end{gathered}
$$

The advantage of this choice of basis rays is that the roportions of the basis rays present in a skew ray are equal to the two object coordinates, $G$ and $H$, for the $x$ and $y$ chief rays, and the two Cartesian stop coordinates, $x_{e}$ and $y_{e}$ for the $x$ and $y$ narginal rays. Thus for the ray from object point (C.H) through stop iocation ( $\mathrm{x}_{\mathrm{e}}, \mathrm{Y}_{\mathrm{e}}$ ), the ray intercept at the $\mathrm{q}^{\text {th }}$ surface is.

$$
\left(x_{q} \cdot y_{q}\right)=\left(G y_{c, q}+x_{e} y_{m, q}, H y_{c, q}+y_{e} y_{m, q}\right) .
$$

Likewise, the ray slowe after the $q^{\text {th }}$ interface is.

$$
\left(u_{x, q}, u_{y, q}\right)=\left(G u_{c, q}+x_{e} u_{m, q}, H u_{c, q}+y_{e} u_{m, q}\right)
$$

## Angle and Plane of Incidence for Radially Symmetric Systems

The angle of incidence, $i_{q}$. for a skew ray must be calculated using the Pythagorean theorem because $\mathrm{i}_{\mathrm{q}}$ has both x and y components:

$$
\begin{aligned}
& i_{x, q}=G i_{c, x, q}+x_{e} i_{m, x, q}=G i_{c, q}+x_{e} i_{m, q} \\
& i_{y, q}=H i_{c, y, q}+y_{e} i_{m, y, q}=H i_{c, q}+y_{e} i_{m, q} \\
& i_{q}=\sqrt{i_{X, q}^{2}+i_{X, q}^{2}}
\end{aligned}
$$

$$
=\sqrt{\left(G^{2}+H^{2}\right) i_{c, q}^{2}+2\left(G x_{e}+H y_{e}\right) i_{c, q} i_{m, q}+\left(x_{s}^{2}+y_{s}^{2}\right) i_{m, q}^{2}} .
$$

For systems radially symmetric about the optical axis, the object can be restricted to the y -axis without loss of generality. Then $\mathrm{G}=0$ and the angle of incidence simplifies to,
or in polar pupil coordinates,

$$
i_{q}=\sqrt{\mathrm{H}^{2} \mathrm{i}_{\mathrm{c}, \mathrm{q}}^{2}+2 \mathrm{H} \rho \cos \phi \mathrm{i}_{\mathrm{c}, q^{1}}{ }_{\mathrm{m}, \mathrm{q}}+\rho^{2 i_{m, q}^{2}}} .
$$

To express the polarization matrix for a ray at an interface, the orientation of the $s$ and p planes with respect to the global system coordinates are required. The p plane is the plane of incidence, that plane which contains the ray and the surface normal at the intersection point. The orientation of the plane of incidence, $\theta$, measured counterclockwise from the $y$ axis, is,

$$
\tan \theta=\frac{i_{x}}{i_{y}} .
$$

If $\mathrm{i}_{\mathrm{x}}=0$, then the ray is directly aoove or below the normal; the plane of incidence then intersects the $\mathrm{x}-\mathrm{y}$ plane in a vertical line. The orientation of the plane of incidence is:

$$
\begin{aligned}
& \sin \theta=\frac{i_{x}}{\sqrt{\left(i_{x}^{2}+i_{y}^{2}\right)}}=\frac{x_{e} i_{m}}{|i|}=\frac{\rho \sin \phi i_{m}}{|i|} \\
& \cos \theta=\frac{i_{y}}{\sqrt{\left(i_{x}^{2}+i_{y}^{2}\right)}}=\frac{H i_{c}+\rho \cos \phi i_{m}}{|i|}
\end{aligned}
$$

Figure 11 shows the paraxial angle and plane of incidence for three field angles. The magnitude of the angle of incidence is represented by the length of the lines. The orientation of the plane of incidence corresponds to the orientation of the lines. Note that off axis. the pattern is just a shifted version of the on axis pattern.

Expressions for $\sin 2 \theta$ and $\cos 2 \theta$ will be required; these relations are:

$$
\begin{gathered}
\sin 2 \theta=2 \sin \theta \cos \theta=\frac{-2 \rho \sin \phi \mathrm{i}_{\mathrm{m}}\left(\mathrm{H} \mathrm{i}_{\mathrm{c}}+\rho \cos \phi \mathrm{i}_{\mathrm{m}}\right)}{\mathrm{i}^{2}} \\
=\frac{-2 \mathrm{H} \rho \sin \theta \mathrm{i}_{\mathrm{c}^{\mathrm{i}} \mathrm{~m}}-\rho^{2} \sin 2 \phi}{\mathrm{i}^{2}} . \\
\left.\cos 2 \theta=\cos ^{2} \theta-\sin ^{2} \theta=\frac{\left(H \mathrm{i}_{\mathrm{c}}+\rho \cos \phi \mathrm{i}_{\mathrm{m}}\right)^{2}-\rho^{2} \sin ^{2} \phi \mathrm{i}_{\mathrm{m}}}{\mathrm{i}^{2}}\right) \\
=\frac{\mathrm{H}^{2} \mathrm{i}_{\mathrm{c}}^{2}+2 \overline{\mathrm{H}} \rho \cos \phi \mathrm{i}_{\mathrm{c}^{2}} \mathrm{i}_{\mathrm{m}}+\rho^{2} \cos 2 \phi \mathrm{i}_{\mathrm{m}}^{2}}{\mathrm{i}^{2}} .
\end{gathered}
$$

## Paraxial Relations for the Angle of Incidence

This section derives the paraxial relations for the meridienal and chief ray angles of incidence from a paraxial ray trace or y y-bar diagram.

Consider an optical system with Lagrange invariant $\Gamma$.

$$
\Gamma=n\left[y u_{c}-\bar{y} u_{m}\right]=n y_{e} u_{e, c}=-n \bar{y}_{\mathrm{l}} u_{\mathrm{I}, \mathrm{~m}} .
$$

Here, n is the refractive index, y and $\overline{\mathbf{y}}$ are the paraxial marginal and chief ray heights and $u_{m}$ and $u_{c}$ are the paraxial marginal and chief ray angles, all in an arbitrary meridional plane. The two final equalities are the Lagrange invariant evaluated at a pupil (with pupil semidiameter $y_{e}$ ) or at an image [with image semidiameter $\bar{y}_{I}$ ]. Semidiameter is used since the term radius is ambiguous, referring both to the half the diameter of the beam of light and to the
radius of curvature. The word radius will be reserved for radius of curvature.
Consider light reflecting (or refracting) at an interiace. Figure 23 shows the general case in the $y$ y-bar diagram format. Let the previous interiace have marginal and chief ray heights of $\mathbf{y}_{\mathrm{d}}$ and $\overline{\mathbf{y}}_{\mathrm{d}}$. On the $\mathrm{y} y$-bar diagram the previous interface is located along line $d$. The following interface has ray heights of $y_{f}$ and $\bar{Y}_{f}$ and is located on line $f$. On the $y$ y-bar diagram the slopes $s$ and $s^{\prime}$ of lines $d$ and f are:

$$
s=\frac{y-y_{d}}{\bar{y}-\bar{y}_{d}} \text {, and, } \quad s^{\prime}=\frac{y-y_{f}}{\bar{y}-\bar{y}_{f}} .
$$

$s$ and $s^{\prime}$ are the ratios of the marginal ray angle to chief ray angle for these portions of "optical space."

$$
s=\frac{u_{m}}{u_{c}}
$$

The semidiameters of the entrance pupil and image in the optical spaces before and after the interface are $y_{e}, y_{e}^{\prime}, \bar{y}_{1}$ and $\bar{y}_{\mathrm{I}}$. Lines $d$ and f have the equations:

$$
\begin{aligned}
y & =y_{e}+s \bar{y}=s\left[y_{I}-\bar{y}\right] . \\
y^{\prime} & =y_{e}^{\prime}+s^{\prime} \bar{y}=s^{\prime}\left[y_{I}^{\prime}-\bar{y}\right] .
\end{aligned}
$$

The chief and marginal ray angles before and after the interface are:

$$
\begin{aligned}
& u_{m}=\frac{-\Gamma}{1 y_{I}}=\frac{-s \Gamma}{n[y-s \bar{y}]} \\
& u_{c}=\frac{\Gamma}{n y_{e}}=\frac{\Gamma}{n[y-s \bar{y}]}
\end{aligned}
$$



Figure 23 Y-Y Bar Parameters for an Interface
The point $[\overline{\mathrm{y}}, \mathrm{y}]$ represents a particular interface. $\left[\overline{\mathrm{y}}_{\mathrm{d}}, \mathrm{y}_{\mathrm{d}}\right]$ and $\left[\overline{\mathrm{y}}_{\mathrm{f}}, \mathrm{y}_{\mathrm{f}}\right]$ are the preceding and following interfaces, on lines $d$ and $f . \bar{y}_{I}$ and $y_{E}$ are the image and pupil located in the optical space preceeding the interface. $\overline{\mathrm{y}}_{\mathrm{I}}$ and $\mathrm{y}_{\mathrm{E}}^{\prime}$ are the image and pupil following the interface.

$$
\begin{aligned}
& u_{m}^{\prime}=\frac{-\Gamma}{n^{\prime} y_{I}^{\prime}}=\frac{-s^{\prime} \Gamma}{n^{\prime}\left[y-s^{\prime} \bar{y}\right]} \\
& u_{c}^{\prime}=\frac{\Gamma}{n^{\prime} y_{e}^{\prime}}=\frac{\Gamma}{n^{\prime}\left[y-s^{\prime} \bar{y}\right]}
\end{aligned}
$$

## Refraction

For a refracting interface with refractive indices $n$ and $n^{\prime}$ before and after the interface, the paraxial law of refraction the first order approximation to Snell's law) is

$$
n i=n^{\prime} i^{\prime}, \text { or, } i^{\prime}=\frac{n}{n^{\prime}} i=N i
$$

If $s_{n}$ is the angle of the surface normal at a ray intercept, then the relation between ray angles and angles of incidence

$$
s_{n}=u-i=u^{\prime}-i^{\prime}
$$

leads to the equation for angle of incidence

$$
\mathrm{i}=\frac{\mathbf{u}^{\prime}-\mathrm{u}}{\mathrm{~N}-1}
$$

The chief and meridional angles of incidence as functions of the $y$ y-bar diagram parameters are:

$$
\begin{aligned}
& i_{m}=\frac{-\Gamma}{N-1}\left[\frac{s^{\prime}}{n^{\prime}\left(y-s^{\prime} \bar{y}\right)}-\frac{s}{n(y-s \bar{y})}\right]=\frac{-\Gamma}{2}\left[\frac{1}{n^{\prime} y_{I}^{\prime}}-\frac{1}{n y_{I}}\right] \\
& i_{c}=\frac{\Gamma}{N-1}\left[\frac{1}{n^{\prime}\left[y-s^{\prime} \bar{y}\right]}-\frac{1}{n[y-s \bar{y})}\right]=\frac{\Gamma}{2}\left[\frac{1}{\bar{n}^{\prime} y_{e}^{\prime}}-\frac{1}{n y_{e}}\right]
\end{aligned}
$$

## Reflection

Consider a reflecting interface immersed in a transmissive medium of refractive index $n$. Reflection can be treated by considering the medium as having a refractive index following the interface, $n^{\prime}$, of

$$
\mathbf{n}^{\prime}=-\mathbf{n} .
$$

With this substitution, the paraxial law of refraction correctly reproduces the law of reflection.

$$
\mathrm{n} \mathrm{i}=\mathrm{n}^{\prime} \mathrm{i}^{\prime} \text {, yields, } \mathrm{i}^{\prime}=-\mathrm{i} .
$$

The relationship between the angle of incidence and the incident and reflected ray angles is

$$
u^{\prime}=u-2 i .
$$

Thus the angle of incidence is

$$
i=\frac{u-u^{\prime}}{2}
$$

The marginal and chief ray angles of incidence, $i_{m}$ and $i_{c}$, expressed in terms of $y$ y-bar diagram parameters are:

$$
\begin{aligned}
& i_{m}=\frac{-\Gamma}{2 n}\left[\frac{s^{\prime}}{y-s^{\prime} \bar{y}}+\frac{s}{y-s \bar{y}}\right]=\frac{-\Gamma}{2 n}\left[\frac{1}{y_{I}^{\prime}}+\frac{1}{y_{I}}\right] . \\
& i_{c}=\frac{\Gamma}{2 n}\left[\frac{1}{y-s^{\prime} \bar{y}}+\frac{1}{y-s \bar{y}}\right]=\frac{\Gamma}{2 n}\left[\frac{1}{y_{e}^{\prime}}+\frac{1}{y_{e}}\right] .
\end{aligned}
$$

## APPENDIX E

## SELECIED BIBLIOGRAPHY

Aggarwal, G. S., "Coherence matrices of light beams at a dielectric interface and the Goos-Hanchen effect." J. Opt. Soc. Am. 65, 1475 (1975). This paper contains a coherency matrix treatment of dieiectric interfaces and includes the degree of polarization.

Azzam. R. M. A., "Direct Relation between Fresnel's interface reflection coefficients for the parallel and perpendicular polarizations," J. Opt. Soc. Am. 69, 1007 (1979). The foliowing reiation between rs and ip is derived:ip $=$ is ( $\mathrm{rs}-2$ $\cos \mathrm{i}) /(1-\mathrm{rs} \cos 2 \mathrm{i})$.

Azzam, R. M. A., "Extinction of the $p$ and $s$ polarizations of a wave on reflection at the same angle from a transparent film on an absorbing substrate: applications to parallel-mirror crossed polarizers and a novel integrated polarimeter," J. Opt. Soc. Am. A. 2, 189 (1985).

Azzam, R. M. A., and N. M.Bashara, Elipsnmetry and Polarized Light (North Holland. Amsterdam. 1977). This book treats in detail the propagation of polarized light through systems of planar interfaces, ellipsometric instrumentation, and the theory and analysis of ellipsometric measurements.

Azzam. R. M. A., and N. M. Bashara. "Polarization Transfer Function of an Optical System as a Bilinear Transformation." J. Opt. Soc. Am. 62. 222 (1972). The mathematics behind the Jones matrix are reformulated to study polarization transfer functions. Let $\chi=E y / E x$, the ratio of complex intensities for polarization states. This $\chi$. defined on the complex plane, can be considered as an alternative to the Poincare sphere for describing coherent polarization states. States of constant azimuth and states of constant ellipticity are both circles in the complex plane. The relation between the incident state, $x$, and the transmitted state, $\xi$, is given by the bilinear transformation upon the Jones matrix elements, $\left.\xi=\mathrm{J}_{22} \chi+\mathrm{J}_{21}\right) /\left(\mathrm{J}_{12} \chi+\mathrm{J}_{11}\right)$. Several interesting results on polarization transfer functions are then derived, such as circles in the incident state space must remain circles in the transmitted state space, and that two eigenpolarizations with different azimuths always exist. Examples applied to ellipsometry are given. This material is also in chapter 2 of their book.

Azzam, R. M. A., and M. E. R. Khan, "Complex reflection coefficients for the parallel and perpendicular polarizations of a film-substrate system." Appl. Opt. 22, 253 (1983). This is a study in various functions of the amplitude transmittance relations for single layer films. This paper contains many fascinating polarization graphs.

Azzam, R. M. A. and K. Javily, "Antireflecting and polarizing transparent bilayer coatings on absorbing substrates at oblique incidence". Appl.Opt 24, 519 (1985) A theory of two layer coatings at oblique incidence is derived. The air $-\mathrm{Si}_{3} \mathrm{~N}_{4}-\mathrm{SiO}_{2}-\mathrm{Si}$ system is used as an example.

Azzam, R. M. A. and B. E. Perilloux, "Constraint on the optical constants of a filmsubstrate system for operation as an external-reflection retarder at a given angle of incidence", Appl.Opt 24, 1171 (1985). This paper treats single layer coatings as retarders and contains lots of graphs.

Azzam, R. M. A. and T. F. Thonn, "Pseudo-Brewster and second-Brewster angles of an absorbing substrate covered by a transparent thin film". Appl.Opt 22, 4155 (1983). Numerical studies of the single layer system were performed with air-SiO2-Si at 633 nm as an example.

Bates, H. E., "Optical Pulse Compression Using Polarizing Techniques", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 90 (1981).

Batten, C. E., "Spectral optical constants of soots from polarized angular reflectance measurements". Appl.Opt. 24, 1193 (1985). A visible/IR polarization ratio spectroreflectometer was used to characterize soot.

Baur, T. G., "Optical polarimeters for solar research", Opt.Eng. 20, 2 (1981).
Bennett, J. M. and H. E. Bennett. "Polarization", Chapt. 10 in Handbook of Optics, edited by W.G.Driscoll and W.Vaughan, McGraw Hill, New York (1978).

Bergmann, E. E., "Polarization-Independent beam splitter", Rev.Sci.Instrum. 54, 579 (1983). A beam splitter was constructed from a thin plate of suitably orientated, uncoated quartz. The thickness is critical.

Blau, H. H., Jr., E. L. Gray, and G. M. B. Bourioius, "Reflection and Polarization Properties of Powder Materials", Appl. Opt. 6, 1899 (1967).

Boissoles, J., and E. Leroux, "Perte de Reflexion en Lumiere Polarisee sur une Surface Metallique Rugueuse au Voisinage de la Frequence du Plasmon de Surface", J. Optics 11, 251 (1980).

Bolt, L., "Hydrolytic Stability of Polyvinyl Alcohol (PVA) Iodine Polarizers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 22 (1981).

Born, M., and E. Wolf, Principles of Optics (Pergamon Press, Oxford, 1975). Chapter 1 treats the Fresnel equations and stratified dielectric media. The optics of metals and crystals are also covered.

Breckenridge, J. B., "Polarization Properties of a Grating Spectrograph", Appl. Opt 10. 286 (1971). For a spectrograph on the McMath solar telescope, the ratio of the $s$ and $p$ intensity transmission coefficients was measured for six diffraction orders. The ratio varied from 0.3 to 20. It is shown how this may cause large photometric and radiometric errors. The Woods anomalies are described in detail.

Brunton, J. H.. "Polarization of the Light Output from a Ruby Optical Maser", Appl. Opt. 3, 1241 (1964).

Chan, C. C., and Tamir, T., "Angular shift of a Gaussian beam reflected near the Brewster angle", Opt. Let. 10, 378 (1985).

Clarke. D., and J.F.Grainger, Polarized Light and Optical Measurement (Pergamon Press. 1971). Polarized light, polarizers, polarization effects at interfaces. the measurement of polarization and the use of polarized light in optical instruments are treated.

Curran, P.J., "A Photographic Method for the Recording of Polarized Visible Light for Soil Surface Moisture Indications". Rem. Sens. Environ. 7, 305 (1978).

Curran, P. J., "The Use of Polarized Panchromatic and False-Color Infrared Film in the Monitoring of Soil Surface Moisture", Rem. Sens. Environ. 8, 249 (1979).

Curran. P. J., "The Relationship Between Polarized Visible Light and Vegetation Amount", Rem. Sens. Environ. 11, 87 (1981).

Curran, P. J., "Polarized Visible Light as an Aid to Vegetation Classification", Rem. Sens. of Envoirn. 12, 491 (1982). lncreased discrimination in remote sensing is possible using polarization data as well as spectral data.

Egan. W. G., "Optical Stokes Parameters for Farm Crop Identification". Rem. Sens. Environ. 1. 165 (1970).

Egan. W. G.. J. Grusauskas, and H. B. Hallock. "Optical Depolarization Properties of Surfaces Illuminated by Coherent Light". Appl. Opt. 7. 1529 (1968).

Fainman, Y., and J. Shamir, "Polarization of nonplanar wave fronts", Appl. Opt 23. 3188 (1984). Lineariy polarized light traversing a polarizer is treated in a diffraction formalism to include the effects of the angular spectrum of the light at a planar interface. Both an ordinary polarizer and Brewster angle reflection polarizer are treated as examples. A picture of the four leaf clover pattern produced by a spherical wave at a polarizer is included.

Gasparian, G. A., "Fiber Optic Reflection Coupler for Bidirectional Transmission, in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 43 (1981).

Gehrels. T.. "Wavelength Dependence of the Polarization of the Sunlit Sky", J. Opt. Soc. Am. 52, 1164 (1962). This paper treats the effect of multiple scattering, molecular anisotropy, reflection from the ground and aerosols in the daytime polarization data.

Gerrard. A.. and J. M. Burch. Introduction to Matrix Methods in Optics (John Wiiey and Sons, London. (1975). Chapter 4 is a detailed treatment of the Jones and Mueller calculus with worked examples and an emphasis on the photoelastic effect. The measurement of the Jones and Mueller matrix elements is also treated.

Goguen. J. D., and W. M. Sinton. "Characterization of Io's Volcanic Activity by Infrared Polarimetry", Science 230. 65 (1985). The number and position of large volcanos as well as their indices of refraction were determined by polarimetry.

Goodman, J. W.. Statistical Optics (John Wiley \& Sons, New York, 1985) . Chapter four ireats the statistical foundations of polarized light, covering the coherency matrix and the polarization (Jones) matrix. The probability density function is derived for the intensity of light of different degrees of polarization.

Harris Jr., J. O., Jr., J. D. Cyrus and G. R. Laguna, "Minimum/Constant Voltage Lead Lanthanum Zirconate Titanate (PLZT) Electro-Optic Shutiers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 53 (1981).

Hauge. P. S., "Techniques of Measurement of the Polarization-Altering Properties of Linear Optical Systems", in Optical Polarimetry, Proc. SPIE 112, 2 (1977).

Hecht. E., and A.Zajac, Optics, Addison-Wesley, Reading, MA. (1974). Chapter 8 is an introduction to polarization with a description of the Iones and Mueller matrices.

Howell. B. J.. "Measurement of the polarization effects of an instrument using partially polarized light". Appl. Opt. 18, 809 (1979) The accuracy of a radiometer is adversely affected by scene polarization if the receiving optical system is sensitive to polarization. The instrument Mueller matrix can be measured and used to determine the maximum polarization sensitivity.

Hsien-Yu Hsu and Yung-Kang Liang, "A Generalized Intensity Formula for a System of Retardation Plates", J. Opt. Soc. Am. 31, 488 (1941).

Jacobs, S. D., "Liquid Crystals as Large Aperture Waveplates and Circular Polarizers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 98 (1981).

Jenkins, F. A., and H. E. White, Fundamentals of Optics, (McGraw-Hill, New York, 1976). An extended discussion of polarization phenomena is contained including a noteworthy description of metallic reflection.

Jones, R. C., "A New Calculus for the Treatment of Optical Systems: I. Description and Discussion of the Calculus", J. Opt. Soc. Am. 31, 488 (1941). Reprinted in Swindell, Polarized Light. 187 (1975). This is the first of the series of papers where the Jones matrix was introduced and developed. It is proved that the matrix representing any combination of ideal retarders and rotators is unitary.

Jones, R. C., "A Now Calculus for the Treatment of Optical Systems: II. Proof of Three General Equivalence Theorems", J. Opt. Soc. Am. 31, 493 (1941). Reprinted in Swindell. Polarized Light. 193 (1975). The following theorems regarding polarizers are proved using the Jones calculus:

1. Any system of only retardation plates and rotators is optically equivalent to a single retarder and rotator,
2. Any number of partial polarizers and rotators is equivalent to a single partial polarizer and rotator, and
3. Any number of retarders, partial polarizers and rotators is equivalent to a system of four elements: two retardation plates, one partial polarizer and one rotatūi.

Jones, R. C.. "A New Calculus for the Treatment of Optical Systems: III. The Sohncke Theory of Optical Activity". J. Opt. Soc. Am. 31, 500 (1941). Reprinted in Swindell, Polarized Light. 200 (1975) This paper treats systems containing a large number of identical retarders, each of which is rotated by an identical angle to the previous retarder.

Jones, R. C., "A New Calculus for the Treatment of Optical Systems: IV.". J. Opt. Soc. Am. 32, 486 (1942). Reprinted in Swindell, Polarized L,ight. 204 (1975). The Jones calcuius is generalized to representations of the Jones vector where the basis vectors are not Cartesian (orthogonal linear polarized states), but are any two linearly independent polarization states. Also, the paper points out that for a system Jones matrix, M, it is a sufficient condition that the eigenvectors be orthogonal if M is unitary or Mermeiian. A necessary condition is that the absolute values of the off diagonal matrix elements are equal, abs(M12) $=\operatorname{abs}($ M21 $)$. Or if $M^{\prime}$ is the Hermitian adjoint of $M$, the necessary and sufficient condition may be stated, that $M$ commute with its Hermetian conjugate. $\mathrm{MM}^{\prime}=\mathrm{M}^{\prime} \mathrm{M}$. It is also shown how the Jones calculus can handle partially polarized light, and proceeds to derive the two polarization states which have maximum and minimum transmittance through a system. The Jones matrix for circular dichroism is derived. The matrix for a monoclinic crystal is derived from a model of alternating
partial polarizers and retarders at an angle to each other in the limit of an infinite number of infinitesimal polarizing elements.

Jones, R. C., "A New Calculus for the Treatment of Optical Systems: V. A More General Formulation, and Description of Another Calculus", J. Opt. Soc. Am. 37, 107 (1947). Reprinted in Swindell, Polarized Light. 212 (1975). It is explained that the Jones calculus is unable to treat systems which depolarize light, but that "the more powerful calculus" of Hans Mueller will.

Jones, R. C., "A New Calculus for the Treatment of Optical Systems: VI. Experimental Determination of the Matrix". J. Opt. Soc. Am. 37, 110 (1947). Reprinted in Swindell. Polarized Light. 216 (1975). A sequence of measurements is described to determine the elements of the Jones matrix of a piece of transparent or crystalline material. When magnetic fields are present, the device may not have the same Jones matrix in both directions and procedures are included for this case.

Jones, R. C., "A New Calculus for the Treatment of Optical Systems: VII. Properties of the N-Matrices". J. Opt. Soc. Am. 38, 671 (1948). Reprinted in Swindell, Polarized Light, 219 (1975). A differential equation is solved for the Jones matrix propagating through a polarizing crystal. The solutions describe the Jones matrix for an infinitesimal length of crystal. For the standard types of polarizing materials, linear and circular polarizers and retarders, these solutions are the Pauli spin matrices, which Jones labels N matrices. Many relationships are given for Jones matrices evolving through systems culminating in a "lamellar representation of a crystal". Uniformly and arbitrarily twisted crystal are also treated.

Jones. R. C., "A New Calculus for the Treatment of Optical Systems, VIII. Electromagnetic Theory", J. Opt. Soc. Am. 37, 107 (1947). Reprinted in Swindell, Polarized Light, 234 (1975). The N matrices of paper VII are related to the dielectric and gyration tensors of crystal optics which describe the anisotropy of the index of refraction and optical activity.

Jones, R. C., "A Comedy of Errors", in Swindell, Polarized Light, 240 (1975). This article clarifies the usage of the gyration tensor in the Jones calculus.

Jones, R. V., and J. C. S. Richards. Proc. Roy. Soc. A225, 122 (London. 1954). This data shows that unpolarized light of 1 micrometer wavelength incident on a steel slit 20 micrometers wide becomes polarized 30 upon passing through the slit.

Kemp. J. C., "Photoelastic-Modulator Polarimeters in Astronomy", in Polarizers and Applications, ed. G. B. Traparii, Proc. SPIE 307, 83 (1981).

Kenyon, R. V., and B. K. Lichtenberg. "Measurement of Ocular Counterrolling (OCR) by Polarized Light", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 79 (1981).

Keussler, Von V.v., and P. Manogg, "Uber die Emission polaristierten Lichtes durch gluhende Metalloberflachen und die dabei vorhandene raumliche Intensitatsverteilung", Optik 17, 602 (1960).

Knowlden, R. E., Wavefront Errors Produced by Multilayer Thin Film Optical Coatings, Dissertation, Univ. of Arizona (1981). This dissertation treats several issues related to the small aberrations that multilayer enhanced reflective coatings introduce into laser systems. Angle of incidence effects due to dielectric reflective coatings are analyzed for a parabolic mirror and an axicon including figures which depict the polarization aberration of these systems. Problems of maintaining and measuring coatings accurately over large areas are analyzed using the technique of two wavelength interferometry. A method for designing and optimizing coatings so that the thickness can be accurately measured is developed.

Kostyla, R. J., and K. J. Abcunas, "Effect of Boration Upon the Stability of Polyvinyl Alcohol (PVA) Iodine Polarizers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 5 (1981).

Kothiyal, M. P., and C. Delisle, "Rotating analyzer heterodyne interferometer: error analysis", Appl. Opt 24, 2288 (1985). The error in phase measurement in an interferometer due to leaky polarization beamsplitter or polarizers is treated. Methods for the use of Jones matrices to analyze polarizing interferometers are developed.

Kubo, H., and R. Nagata, "Stokes Parameters Representation of the Light Propagation Equations in Inhomogeneous Anisotropic, Optically Active Media", Optics Comm. 34, 306 (1980).

Kuga, Y., L. Tsang, and A. Ishimaru, "Depolarization effects of the enhanced retroreflectance from a dense distribution of spherical particles", J. Opt. Soc. Am. A. 2, 616 (1984). This is a polarization study of enhanced backscattering.

Lee, K. O., and Y. P. Neo, "Polarization Irradiance Distribution in the Spectrum of Infrared Grating Instruments", J. Opt. Soc. Am. 61, 273 (1971). This paper contains data on Woods anomalies but few conclusions.

Leonard, D. A., B. Caputo and J. D. Fridman, "Remote Sensing of Subsurface Water Temperature by Raman Polarization Spectroscopy", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 76 (1981).

Lindgren, R. W., and T. D. Tarbell, "Video image processor on the Spacelab 2 Solar Optical Universal Polarimeter (SL2 SOUP)", in Shuttle Pointing of ElectroOptical Experiments, Proc. SPIE 265, 47 (1981).

Lotem, H., and Laor, U., "Low-loss bireflectant (double-reflection) polarization prism", Appl. Opt. 25, 1271 (1986).
Lotsch, H. K. V., "Reflection and Refraction of a Beam of Light at a Plane Interface", J. Opt. Soc. Am. 58, 551 (1968). The Fresnel equations and the Goos-Hansch shift are given an electromagnetic wave treatment which includes a discussion of finite width beams.

Love, J. D., and A. W. Snyder, "Generalized Fresnel's laws for a curved absorbing interface". J. Opt. Soc. Am. 65, 1072 (1975). The optical power transmitted across a curved interface is treated as the sum of two terms, that due to absorbing media when the interface is flat, and that due to curvature of the interface when the media are nonabsorbing.

Marathay, A. S., "Operator Formalism in the Theory of Partial Polarization", J. Opt. Soc. Am. 55, 969 (1965).

Mermelstein, M. D., "All-fiber polarimetric sensor", Appl. Opt. 25, 1256 (1986).
Miles, G. M., "Practical Prism Polarizers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 2.(1981).

Millikan, R. A., Phys.Rev. 3, 177 (1895). This paper may be the origin of the hypothesis that emission polarization is due strictly to the effect of the Fresnel coefficients as light refracts out of the source.

Mueller, H., "The foundations of optics", J. Opt. Soc. Am. 38, 661 (1948).
Okuno, Y., and T. Matsuda, "Efficient technique for numerical solution of diffraction by a sinusoidal grating", J. Opt. Soc. Am. A. 2, 572 (1985).

Perregaux, A. E., "Polarizers for Liquid Crystal Devices: The User's Viewpoint", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 70 (1981).

Petit, R., ed., Electromagnetic Theory of Gratings, (Springer-Verlag, Berlin, 1980).
Parke III, N. G., "Matrix Optics", Ph.D. dissertation, Dept.of Physics, Massachusetts Institute of Technology, (1948).

Roese, J. A., "Advances in Electro-Optic Shutter Stereoscopic Displays", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 46 (1981).

Rosenberg, W. J., and A. M. Title, "Solc Filter Engineering", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 106 (1981).

Russell, E. E., L. A. Watts, S. F. Pellicori, and D. L. Coffeen, "Orbiter Cloud Photopolarimeter for the Pioneer Venus Mission", in Optical Polarimetry, Proc. SPIE 112, 28 (1977).

Saito, K., H. Noda, and T. Namioka, "Polarization Characteristics of Concave Gratings Mounted in a Seya-Namioka Monochromator", Optica Acta 25. 1055 (1978). This paper contains experimental data with conclusions for eighteen gratings.

Samek, N. E. and M. A. Munoz "Light Transmission Characteristics of a High Density Electro-Optical Light Gate Array Based on Lerad Lanthanum Zirconate Titanate (PLZT) Ceramics", in Polarizers and Applications, ed. G. B. Trapani. Proc. SPIE 307, 60 (1981).

Sandus, O., "A Review of Emission Polarization". Appl. Opt 4, 1634 (1965). Emission polarization can be explained by the refraction of all the emitted radiation, in accord with Millikan's hypothesis of 1895. Data and theories of black body polarization effects are collected.

Shurcliff, W. A.. Polarized Light, Production and Use, (Harvard University Press, Cambridge. MA, 1961). This book contains an extended treatment of polarized light, all types of polarizers. polarized light systems, and pciarization calculi. There is an extensive bibliography of earlier material on polarization phenomena.

Shellan, J. B., "Phased-array performance dsgradatinn due to mirror misfigures, piston errors, jitter, and polarization errors", J. Opt. Soc. Am. A. 2, 555 (1985). A very short section asserts that the effect of polarization variations (errors) on the Strehi ratio of a phased ariay optical system is equivalent (to lowest order) to optical path errors in the different channels.

Seeley, J. S., "Simple nonpolarizing high-pass filter". Appl. Opt 24, 742 (1985). Filter designs to reduce polarization differences in edge filters are described.

Seward III. T. P., "Glass Polarizers containing silver", Proc. SPIE 464, 48 (1984). This paper includes a description of polarizers whose dichroic ratio can vary with position due to suitable high intensity irradiation during fabrication.

Simon, J. M., and M. C. Simon, "Elliptic Polarization in Grating Spectrographs", Appl. Opt 12.153 (1973) Light polarized at 45 degrees to the grating rulings was studied in an Ebert spectrograph.

Slocum, R. E., "Evaporative Thin Metal Films as Polarizers", in Polarizers and Applications, ed. G. B. Trapani. Proc. SPIE 307. 25 (1981).

Smith, D. Y.. "Comments on the dispersion relations for the complex refractive index of circularly and elliptically polarized light". J. Opt. Soc. Am. 66, 454 (1976). The Kramers-Kronig relations for circularly polarized modes are discussed with application to circular polarization and circular retardance.

Stewart J. E., and W. S. Gallaway, "Diffraction Anomalies in Grating Spectrometers", Appl. Opt. 1, 421 (1962). Spectrophotometric errors introduced by Woods anomalies are described. The paper ends with a challenge to theorists to explain the results

Stenflo, J. O., "Solar magnetic and velocity-field measurements: new instrument concepts", Appl. Opt. 23, 1267 (1984).

Stenflo, J. O., "Stokes Polarimetry", in Proc. Workshop on Small Magnetic Flux Concentrations in the Solar Photosphere (1985).

Stenflo, J. O., and H. Povel, "Astronomical polarimeter with 2-D detector arrays", Appl. Opt. 24, 3893 (1985).

Stenflo J. O., and H. Povel, "Astronomical polarimetry with two-dimensional detector arrays", Preprint - submitted to Appl. Opt.

Stokes, G. G., "On the Composition and Resolution of Streams of Polarized Light from Different Sources", Trans. Cambridge Phil.Soc., 9, 399 (1852). Reprinted in Swindell, Polarized Light, 124 (1975). Introduction of the "Stokes parameters".

Swindell, W.. "Handedness of Polarization after Metallic Reflection of Linearly Polarized Light", J. Opt. Soc. Am. 61, 212 (1971). A careful treatment of the reflection of polarized light from metals.

Swindell, W., "Handedness of Polarization after Total Reflection of Linearly Polarized Light", J. Opt. Soc. Am. 62, 295 (1972).

Swindell, W., Polarized Light in Optics, (Dowden, Hutchinson \& Ross, Stroudsburg, PA, 1975). This book is a collection of 35 classic papers on polarized light. Particularly noteworthy is the inclusion of R. Jones original series of papers on the Jones calculus.

Theocaris, P. S., and E. E. Gdoutos, Matrix Theory of Photoelasticity, (SpringerVerlag, Berlin, 1979). This book is a careful introduction to polarized light, polarizers and polariscopes with application to two and three dimensional photoelasticity. Interferometric and holographic photoelastic methods and the method of birefringent coatings are included.

Thornton, A. L., "Electro-Optic Shutter Devices Utilizing Lead Lanthanum Zirconate Titanate (PLZT) Ceramic Wafers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 32 (1981).

Title, A. M., "Partial polaroids in birefringent filters", Solar Phys. 38, 523 (1974). Lyot birefringent fiiters are treated as combining light which traversed a network of different time delays. The effects of partial polarizers rather than ideal polarizers on the filter line spread function is developed in a Fourier domain treatment.

Title. A. M., and W. J. Rosenberg. "Achromatic Retardation Plates", in Polarizers and Applications. ed. G. B. Trapani, Proc. SPIE 307, 120 (1981).

Trapani, G. B., and K. J. Abcunas, "Color Contrast in Polarizer/Retarder Optical Systems", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 112 (1981).

Vanderbilt, V. C., L. Grant, and C. T. Daughtry, "Polarization of Light Scattered by Vegetation". Proc. IEEE 73, 1012 (1985). Importance of polarization analysis for remote sensing of plants is stressed with lots of empirical data.

Walraven, R., "Polarization imagery", in Optical Polarimetry, Proc. SPIE 112, 164 (1977).
'Walraven, R., "Polarization imagery". Opt.Eng. 20. 14 (1981).
West, S. C., "Polarimetry with multiple mirror telescopes", Proc. SPIE 628, (1985). The instrumental polarization of the beam combining optics (aluminum coated) in the multiple mirror telescope is calculated using the Mueller calculus to be as high as 8 as the telescope tracks across the sky. The accuracy improves in the infrared. The primary and secondary mirror are neglected as relatively small effects.

Westerveld, W. B., K. Becker, P. W. Zetner, J. J. Corr and J. W. McConkey, "Production and measurement of circular polarization in the VUV". Appl. Opt. 24, 2256 (1985). This is an analysis of reflection methods for producing polarized light. It centains data for gold reflectors.

Whitley, J. B.. R. E. Cuthrell and D. M. Mattox. "Continuously Variable ElectroOptic Filter". in Polarizers and Applications, ed. G. B. Trapani. Proc. SPIE 307. 39 ( 198 i ).

Whitney, C., "Pauli-Algebraic Operators in Polarization Optics". J. Opt. Soc. Am. 61. 1207 (1S71).

Worthing. A. G.. "Deviation from Lambert's Law and Polarization of Light Emitted by Incandescent Tungsten, Tantalum and Molybdenum and Changes in the Optical Constants of Tungsten with Temperature". J. Opt. Soc. Am., 13, 635 (1926).

Yeh. P.. "Generalized Model for Wire Grid Polarizers", in Polarizers and Applications, ed. G. B. Trapani, Proc. SPIE 307, 13 (1981).

## REFERENCES

Azzam, R. M. A., and N. M. Bashara, Ellipsometry and Polarized Light, North Holland, Amsterdam, (1977).

Baur, T. G., L. L. House, and H. K. Hull, "A Spectrum Scanning Stokes Polarimeter", Solar Physics 65, 111 (1980).

Born, M. and E. Wolf, Principles of Optics, Pergamon Press, Oxford, (1975).
Breckenridge, J. B., "Polarization Properties of a Grating Spectrograph", Appl. Opt 10. 286 (1971).

Chipman, R. A., "Polarization Aberrations of Lenses", Proc. SPIE 554, 82 (1986).
Clout. P. N., and W. O. Heddle, "Elimination of Polarization Corrections form Optical Excitation-Function Measurcmenis". J. Opt. Soc. Am. 59, 715 (1969).

Clarke, D. and J. F. Grainger, Polarized Light and Optical Measurement. Pergamon Press, (1971).

Cohen-Tannoudji, C.. B. Diu and F. Laloe, Quantum Mechanics, John Wiley and Fons, New York, (1977).

Fainman, Y., and J. Shamir, "Polarization of nonplanar wave fronts", Appl. Opt 23, 3188 (1984).

Fresnel, A., Oeuvres completes de Fresnel, Vol.1\&2, Ed. H. de Senarmont. E.Verdet and L.Fresnel, (1866), Johnson Reprint Co., New York (1965).

Gerrard, A., and J. M. Burch, Introduction to Matrix Methods in Optics, John Wiley and Sons, London, (1975).

Goodman, J. W., Statistical Optics, John Wiley \& Sons, New York, (1985).
Gottfried, K., Quantum Mechanics, W.A.Benjamin, Reading, MA, (1966).
Hecht. E., and A. Zajac, Optics, Addison-Wesley, Reading, MA, (1974).
Hege, E. K., J. M. Beckers, P. A. Strittmatter, and D. W. McCarthy. "Multiple mirror telescope as a phased array telescope", Appl. Opt. 24, 2565 (1985).

Howell, B. J., "Measurement of the polarization effects of an instrument using partially polarized light", Appl. Opt. 18. 809 (1979).

Inoue. S., and W. L. Hyde. "Studies on depolarization of light at microscope lens surfaces", J. Biophys. Biochem. Cyto. 3, 391 (1957).

Jones, R. C., "A New Calculus for the Treatment of Optical Systems. I. Description and Discussion of the Calculus", J. Opt. Soc. Am. 31, 488 (1941) . Reprinted in Swindell, 187 (1975).

Jones, R. C., "A New Calculus for the Treatment of Optical Systems. II. Proof of Three General Equivalence Theorems", J. Opt. Soc. Am. 31, 493 (1941), Reprinted in Swindell. Polarized Light. 193 (1975).

Jones, R. C., "A New Calculus for the Treatment of Optical Systems. III. The Sohncke Theory of Optical Activity", J. Opt. Soc. Am. 31, 500 (1941), Reprinted in Swindell, Polarized Light, 200 (1975).

Jones, R. C., "A New Calculus for the Treatment of Optical Systems. IV.": J. Opt. Soc. Am.32, 486 (1942). Reprinted in Swindell. Polarized Light. 204 (1975).

Jones, R. C., "A New Calculus for the Treatment of Optical Systems. V. A More General Formulation, and Description of Another Calculus", J. Opt. Soc. Am. 37, 107 (1947), Reprinted in Swindell, Polarized Light. 212 (1975).

Jones, R. C., " $\dot{A}$ New Calculus for the Treatment of Optical Systems. VI. Experimental Determination of the Matrix", J. Opt. Soc. Am. 37, 110 (1947) . Reprinted in Swindell, Polarized Light. 216 (1975).

Jones, R. C., "A. New Calculus for the Treatment of Optical Systems. VII. Properties of the N-Matricies". J. Opt. Soc. Am. 38. 67! (1948) , Reprinted in Swindell, Polarized Light. 219 (1975).

Jones, R. C., "A New Calculus for the Treatment of Optical Systems. VIII. Electromagnetic Theory", J. Opt. Soc. Am. 37. 107 (1947) . Reprinted in Swindell, Polarized Light, 234 (1975).

Kingslake, R., Lens Design Fundamentals, Academic Press, New York (1978).
Knowlden, R. E., Wavefront Errors Produced By Multilayer Thin-film Coatings, Ph.D. dissertation, Optical Sciences Center, Univ. of Arizona, (1981).

Kubota, H. and S. Inoue. "Diffraction Images in the Polarizing Microscope", J. Opt. Soc. Am. 49. 191 (1959).

Macleod, A., Thin Film Optical Filters. Macmillan, New York (1986).

Malus, E. L., "Sur une propriete de la lumiere reflechie", (Concerning a property of reflected light), Mem. Phys. Chim. Soc. D'Arcueil, 2, 143 (1809) , Reprinted in Swindell, Polarized Light, 219 (1975).

McCarthy, D. W., P. A. Strittmatter, E. K. Hege, and F. J. Low, "MMT as a OpticalInfrared Interferometer and Phased Array", Proc. SPIE 332, 57 (1982).

Mueller, H., Informal notes for course 8.26 at Massachusetts Institute of Technology, (1943).

Mueller, H., "The foundations of optics", J. Opt. Soc. Am. 38, 661 (1948).
Parke III, N. G., Matrix Optics, Ph.D. dissertation, Dept.of Physics, Massachusetts Institute of Technology, (1948).

Shellan, J. B., "Phased-array performance degradation due to mirror misfigures, piston errors, jitter, and polarization errors", J. Opt. Soc. Am. A. 2, 555 (1985).

Shurcliff, W. A., Polarized Light, Production and Use, Harvard University Press, Cambridge, MA, (1961).
W. J. Smith, W. J., Modern Optical Engineering, McGraw-Hill, New York (1966).

Stavroudis, O. N., Modular Optical Design, Springer-Verlag, Berlin (1982).
Stenflo, J. O., "The measurement of solar magnetic fields", Rep. Prog. Phys., 41, (1978).

Stewart, J. E., and W. S. Gallaway, "Diffraction Anomalies in Grating Spectrometers", Appl. Opt. 1, 421 (1962).

Stokes, G. G., "On the Composition and Resolution of Streams of Polarized Light from different Sources", Trans. Cambridge Phil. Soc., 9, 399 (1952) , Reprinted in Swindell, Polarized Light, 124 (1975).

Swindell, W., Polarized Light in Optics, Dowden, Hutchinson \& Ross, Stroudsburg, PA, (1975).

Theocaris, P. S., and E. E. Gdoutos, Matrix Theory of Photoelasticity, SpringerVerlag, Berlin, (1979).

West, S. C., "Polarimetry with multiple mirror telescopes", Proc. SPIE 628, 50 (1985).
Woolsey, J. M., and J. W. McConkey, J. Opt. Soc. Am. 58, 1309 (1968).
Yariv, A., and P. Yeh, Crystal Optics, John Wiley and Sons, New York (1984).

