Polarization Critical Optical Systems: Important Effects and Design Techniques

Karlton Crabtree
College of Optical Sciences
The University of Arizona
Opti 521
14. December 2007

Table of Contents

Introduction:

Working with polarization effects can be compared to the oriental board game "igo". The rules can be stated quite simply, but the number surprises in the various combinations is sufficiently large that one could study for a lifetime. As there are no good references currently available on polarization critical system design, this paper is based principally on the author's calculations and experience, with reference to various sources of useful information.

This paper will first address the primary sources of polarization effects. As the language of polarization has been evolving very rapidly for the past two decades, this paper will attempt to define all relevant terms, and note common points of confusion. After the sources of the polarization effects have been discussed, several examples will be presented, applying these to common optical components. Finally, the discussion will move to complete systems, and how to minimize the polarization errors induced.

Common Sources of Polarization Effects:

The most common source of polarization effects in optical systems are: Fresnel Effects, Thin Film Effects, Material Birefringence, and Dichroism (preferential absorption of one polarization state). Scatter is a phenomenon, common in nature, which can either polarize (as in the sky) or depolarize (as in an integrating sphere). Scatter, while used in illumination systems, is beyond the scope of this report and will not be addressed further, as this report concentrates on imaging type systems. Structured elements such as diffractive elements and sub-wavelength structures, also beyond the scope of this report, are not covered despite ever increasing use.

Birefringence is a difference in refractive index between two orthogonal polarization states in a material. These two polarization states are transmitted through the material without change, and are called eigenpolarizations. Today the term birefringence is usually used today when the eigenpolarizations are linear, while the term optical activity is used if the eigenpolarizations are the circular polarizations. Eigenpolarizations of materials are most often linear, but can be any orthogonal pair of elliptical polarization states. Birefringence can be caused by the crystalline structure of the material, a stress field either inherent in or applied to the material, application of an electric or magnetic field, etc. The properties of a device which is birefringent will be the same, regardless of the source of the birefringence. For the remained of this paper, only birefringent materials having linear eigenpolarizations will be discussed.

Birefringence in an optical element leads to different optical path lengths, hence different wavefront aberrations, depending on the incident polarization state. Considering a single ray, the difference in the OPD of the two eigenpolarizations is referred to as the retardance, which is usually given the symbol δ. Retardance is most commonly stated in wavelength dependent terms – either waves, degrees, or radians. It is also correct, though less common, to eliminate the wavelength dependence and express the retardance in nm. The polarization state with the shortest OPD is called the fast axis of the material, while the polarization state with the longest OPD is called the slow axis. Since the OPD of the two eigenpolarizations is different, the incident polarization state can be changed, which is the best known use for the plates of birefringent material commonly called quarter and half wave plates. Waveplates, properly called linear retarders, will be discussed in more detail later.

Dichroism is a preferential absorption of one polarization state. The eigenpolarizations of these materials are usually linear, though there is no reason to suppose circular and elliptically dichroic materials cannot exist. Materials exhibiting linear dichroism have many different physical forms, but can generally be modeled as a series of nano-wires. These nano-wires absorb (convert to heat) electric fields polarized in the same orientation as the wires, while largely transmitting the orthogonal polarization state. Today there are two common measures for dichroic materials. One, most often used for measuring polarizer quality, is the contrast ratio, the ratio of transmittance of two polarizers with their transmittance axis parallel, $T_{parallel}$ divided by the transmittance of the same two polarizer with their

axes crossed $T_{perpendicular}$. $CR = \dfrac{T_{¿}}{T_{perpendicular}}$ A second quantity, called diattenuation, defined as

$D = \dfrac{T_{max} - T_{min}}{T_{max} + T_{min}}$ where T_{max} is the transmittance of the most highly transmitted polarization state and T_{min} is the transmittance of the least transmitted polarization state. Diattenuation is more often used for weakly polarizing elements.

Birefringence and Dichroism are bulk material properties which gradually affect the polarization. Polarization effects also occur abruptly at surfaces. For uncoated interfaces, the Fresnel Coefficients describe the polarization effects which occur at the surface (Jackson is one of the few resources which state the full form of the Fresnel Coefficients). They describe polarization effects at transmitting surfaces, reflecting surfaces, and total internally reflecting (TIR) surfaces.

For uncoated dielectric surfaces in transmission and external reflection, the s & p-polarized transmittances are different creating diattenuation, but there is no retardance. For an average glass, at

moderate angles of incidence, the diattenuation is only a few percent in transmission while rapidly increasing in reflection. As the incident angle reaches Brewster's angle, the reflective diattenuation becomes unity, while the transmitted diattenuation is perhaps 15%. See Figure 1, Figure 2, Figure 3, and Figure 4 for plots of these parameters for an average glass.
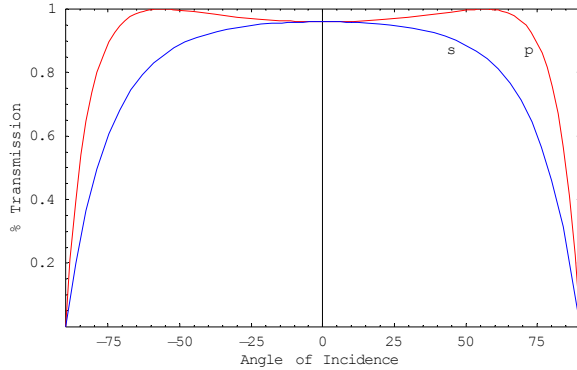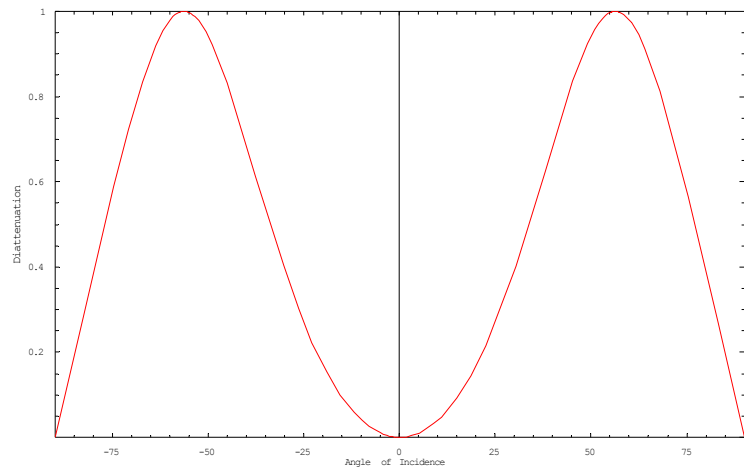

Figure 1: s & p Transmittances at Dielectric Interface
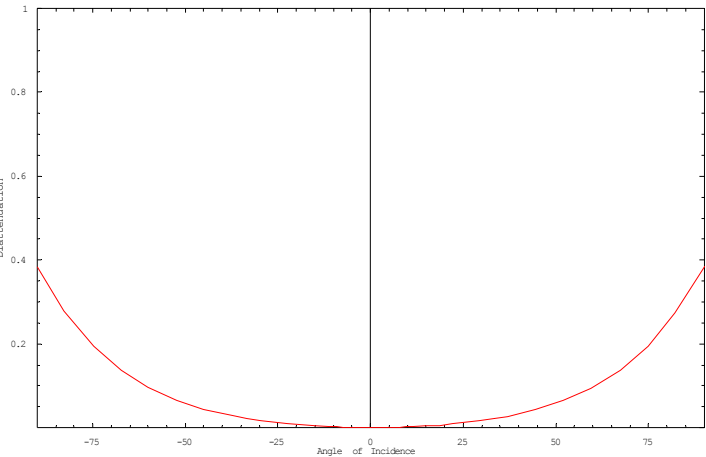
Figure 2: Transmitted Diattenuation at Dielectric Interface
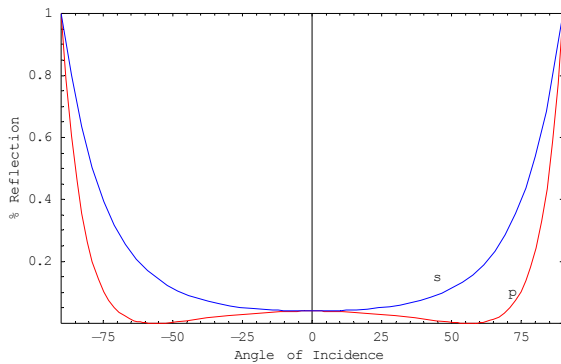

Figure 3: s & p Reflectances at Dielectric Interface

Figure 4: Reflected Diattenuation at Dielectric Interface

At a dielectric interface in internal reflection, 100% of the light is reflected, so there can be no diattenuation. Retardance, however, is significant. For ordinary glass interfaces in TIR there is approximately 1/8th wave of retardance when the angle of incidence is 50°. The retardance decreases to zero as grazing incidence is approached (Figure 5).
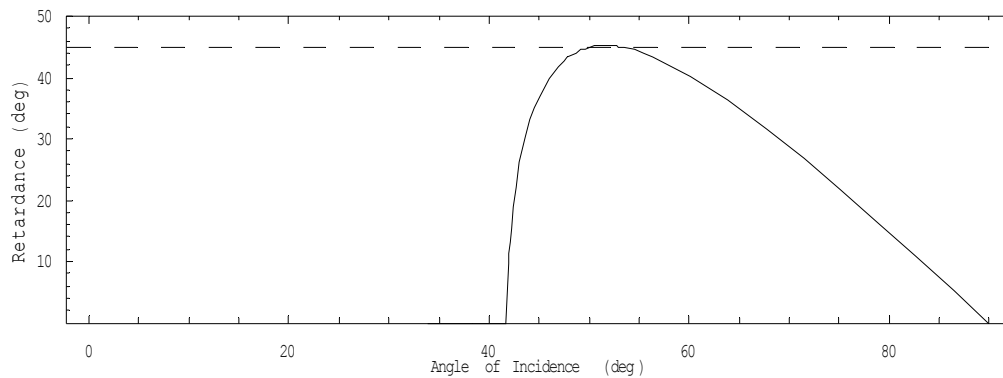
Figure 5: Internal Reflection Retardance of n=1.5 Interface

Thin Film Effects are similar in some ways to Fresnel effects. Both occur suddenly at the surface. Both go to zero at normal incidence. The variation of thin film stacks as a function of both wavelength and angle of incidence can be quite complex. Also, thin films can force the s-polarization of a surface to have greater transmittance than the p-polarization. This drives the diattenuation negative, allowing some polarization correction. The retardance can also be forced negative by clever thin film design.

Polarization of Various Optical Components:

First consider an optical window. Assuming it has uncoated surfaces, it reflects a small fraction of the light incident, while transmitting most of it. If the light is incident at normal incidence, there are no polarization effects. If the window is tilted, there will be diattenuation. In some systems, the polarization state of the beam is predictable, in which case the windows can be aligned at Brewster's angle, and will transmit 100% of the incident light. As a historical note, it was once common to use a reflection from a plate of glass, oriented at Brewster's angle to produce highly polarized light. It is not an efficient method, but the polarization purity can be high. Today, polarizing beam splitters (discussed later) are used. In gas laser tubes, Brewster windows, due to the preferential transmission of one polarization state, cause the laser output to take a particular linear polarization. If the polarization state incident on the window is not predictable, it is generally desirable to place the windows at normal incidence to minimize diattenuation and retardance. Applying a good anti-reflection coating necessarily reduces diattenuation, while introducing retardance from thin film effects. When a spherical wavefront is incident on a window, there are a range of angles of incidence, in a rotationally symmetric pattern. The angles of incidence vary approximately linearly along any particular radial line, making the diattenuation and retardance variation approximately quadratic. This rotationally

symmetric quadratic pattern is, mathematically speaking, quite like the defocus term encountered in scalar aberration theory. As such, we can refer to the diattenuation and retardance patterns as diattenuation defocus, and retardance defocus.

A lens can be considered in much the same way as a window. Again uncoated surfaces will be pure diattenuators, while anti-reflection coated surfaces will have much reduced diattenuation and some retardance. In a planar wavefront, a spherical lens will produce rotationally symmetric, quadratically varying diattenuation and retardance patterns. With a spherical wave incident, two quadratic surfaces are interacting, the wavefront and the lens, producing fourth order interactions. A pattern which is fourth order in radius, and rotationally symmetric is the definition of spherical aberration in scalar aberration theory. Therefore, we can call the diattenuation and retardance patterns with that form diattenuation spherical aberration and retardance spherical aberration respectively.

Spherical mirrors have much the same polarization aberration patterns as a spherical lens. Since mirrors are usually metal coated, they are more likely to simultaneously have significant diattenuation and retardance.

It is undoubtedly obvious to the reader at this point that every aberration form in scalar wavefront aberration theory will also appear in both diattenuation and retardance patterns. The above has only considered the case of zero field. Chipman was the first to write out a polarization aberration theory, utilizing Jones matrices which simultaneously incorporate the effects on the polarization state, the traditional scalar wavefront effects, and the effects of the polarization state on the wavefront.

A nice example of the effects of the polarization state on the wavefront aberration is an F/1.5 Cassegrain telescope (Figure 7) where the mirrors are coated with the coating from MacLeod p.196. This produces a system which is diffraction limited for zero field when unpolarized light is incident, but has half a wave of astigmatism when polarized light is incident. This is due simply to the thin film effects of the surfaces – the phase shift on reflection is quite different for s & p – polarized light, which is to say this coating exhibits significant retardance. In (Figure 6), vertically polarized light is incident, coinciding with the p-polarization plane of the mirror along the vertical axis, and with the s-polarization state of the mirror along the horizontal axis. This is an extreme example – in most optical systems this effect is significantly less than 0.1 waves. Switching this telescope to bare Aluminum coatings would also reduce this effect substantially.
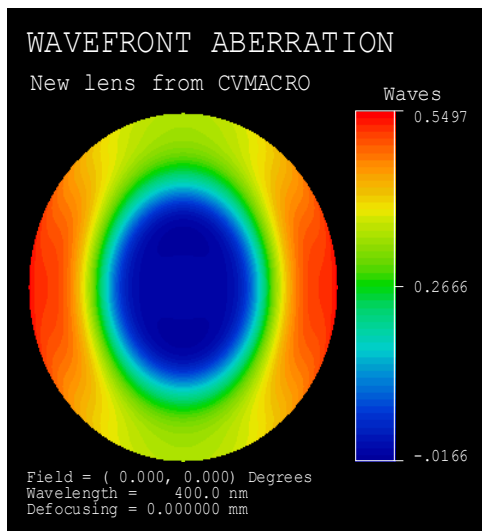
Figure 6: Wavefront Aberration with Vertically Polarized Light Incident
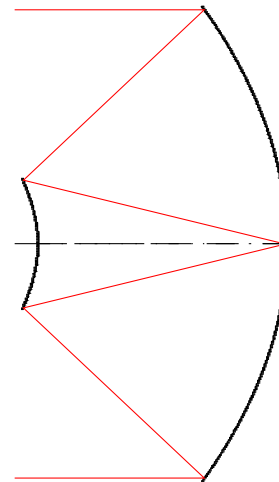
Figure 7: Fast Cassegrain Reflector (F/1.5)

Beam folding elements, such as prisms, fold mirrors, and beam splitters deserve consideration separately. These elements are, in geometrical optics, typically deployed at a 45° angle of incidence to the beam propagation. They will, therefore, have large diattenuation and/or retardance from the Fresnel effects. In metal fold mirrors, both effects are large, while TIR prisms show only retardance. Figure 8 and Figure 9 show the effects of a silver coated mirror at 450 nm. The retardance, from a 30° angle of incidence to a 60° angle of incidence varies from 16° to 70°. This is $1/20^{th}$ of a wave to $1/5^{th}$ of a wave – a huge variation across the pupil (Figure 12, Figure 13). Figure 11 shows how little the incident polarization is changed if it is aligned with the mirror s or p – polarizations. Figure 10 shows the enormous changes which can occur when the incident polarization is oriented at 45° to the mirror s & p – polarization. Changing to other materials can improve this – aluminum for example is almost 10X lower retardance at this wavelength. Also, using two mirrors arranged like a pentaprism, having an angle of incidence of 22.5° each will improve the situation by a factor of two, since the retardance of each reflection is roughly proportional to the square of the angle of incidence, but there are twice as many reflections.
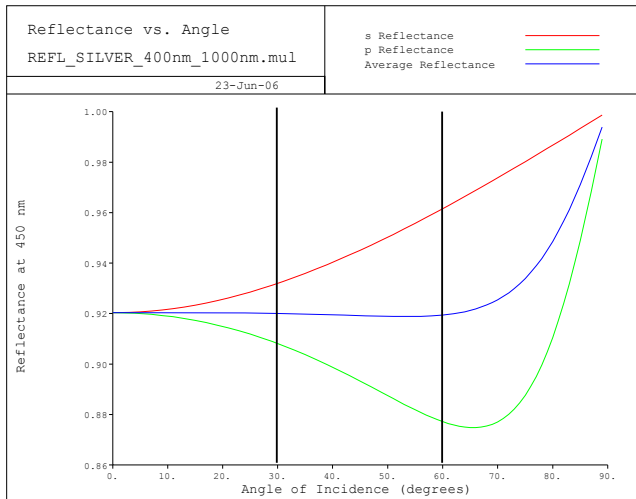
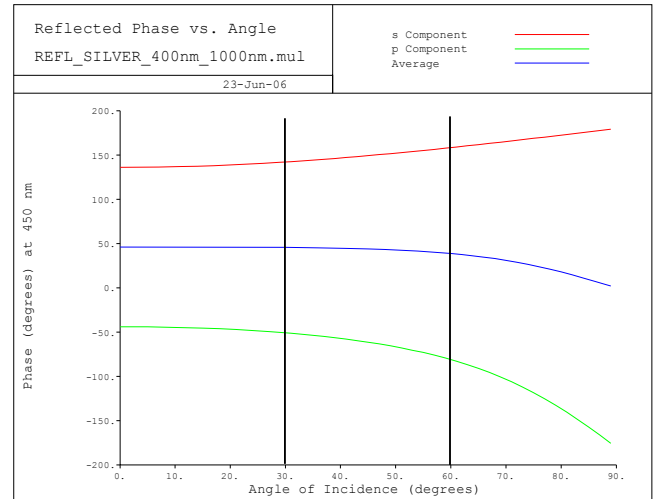*Figure 8: s & p Reflectance for Bare Silver Mirror Coating*

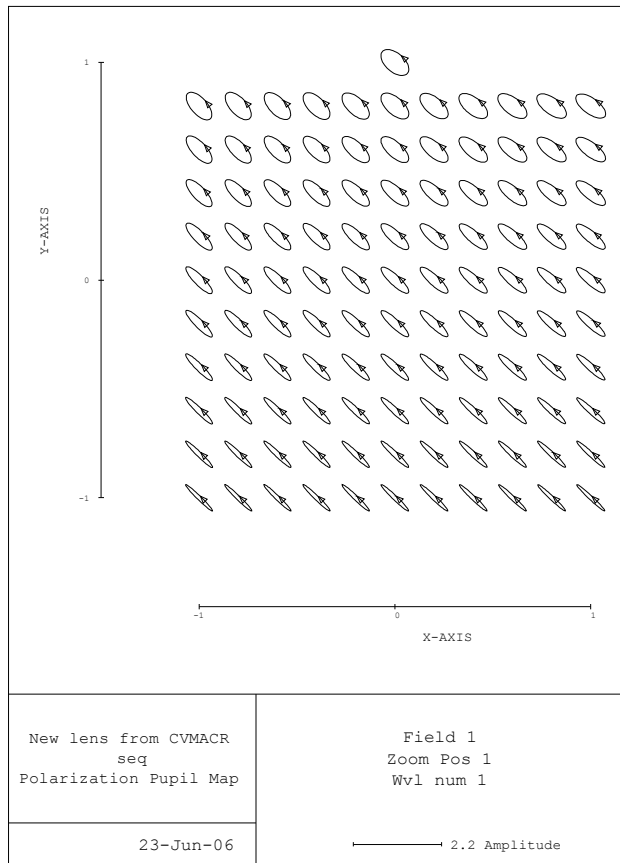*Figure 9: Phase of s & p Polarizations on Bare Silver Mirror*

Figure 10: Exiting Polarization State for Vertical Polarization Incident on Silver Fold Mirror
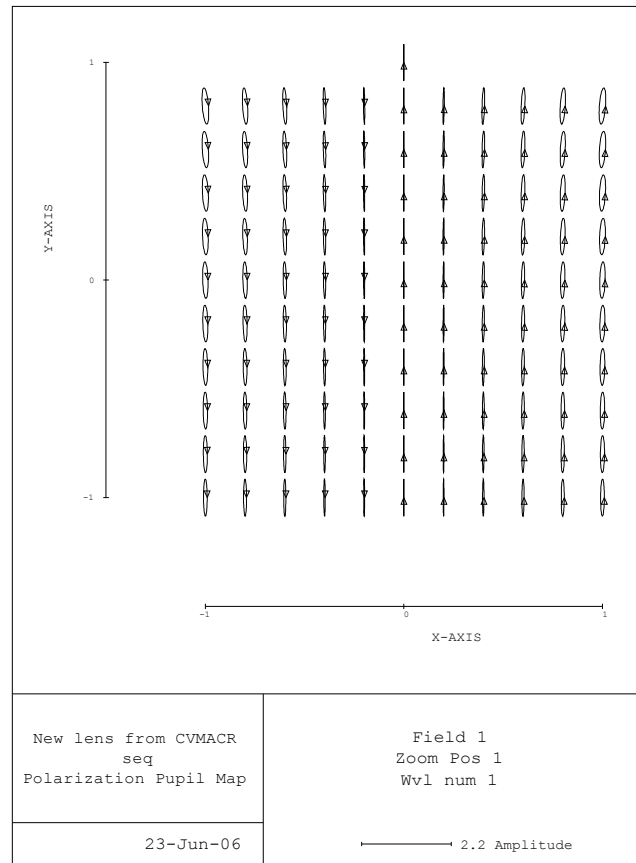

Figure 11: Exiting Polarization State for Vertical Polarization Incident on Silver Fold Mirror
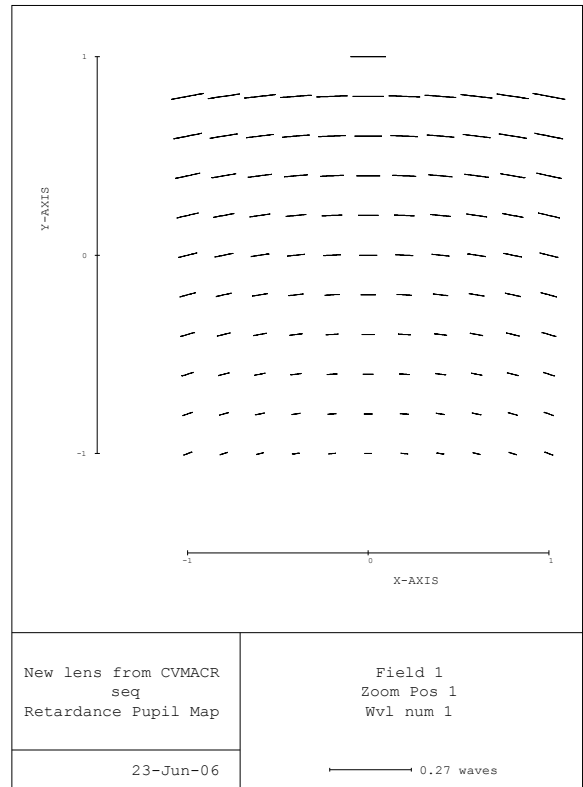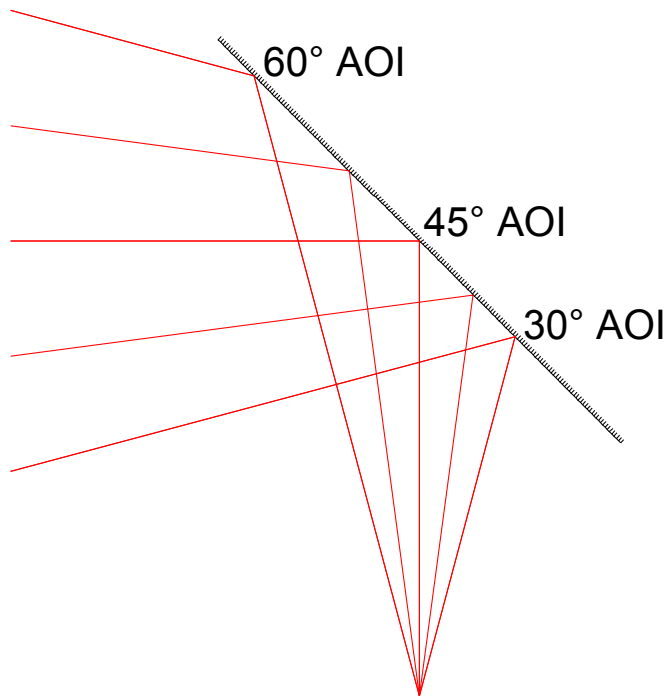
Figure 12: Geometry for Plots of Polarization Due to Silver Fold Mirror

Figure 13: Retardance Aberration of Silver Fold Mirror

A 45° TIR right prism will of course have zero diattenuation, but as shown in Figure 5, located in the first section of this paper, the retardance will be approximately 1/8th wave. Other prisms will suffer a similar retardance at each reflection. These retardances are additive if all reflections are in the same plane, as in the Abbe and Pechan prisms, for example. The Fresnel Rhomb is a rhomboid prism manufactured such that each reflection provides exactly 1/8th wave of retardance, and ¼ wave for the total prism. These are used as highly achromatic quarter wave linear retarders (quarter wave plates). If all reflections do not lie in the same plane, the calculations can become complex, and should be handled on a case-by-case basis.

Beam splitters are a special case. Today, the most common beam splitters use very complicated thin film dielectric stacks. This has resulted in high quality (>99% diattenuating) polarizing beam splitters being readily available. So called Non-Polarizing Beam Splitters (NPBS) do not have equivalent quality as yet. Today, the best NPBS which are commonly available will reliably split an incident beam into two beams having very close to 50% of the incident irradiance. They tend to have 5% or more diattenuation, meaning that when unpolarized light is incident, the resulting beams are 5% polarized, and 20° of retardance, meaning when some polarization state other than the s or p

-polarization is incident, the emerging polarization state is noticeably different from the incident state. The spatial uniformity also tends to be poor. In addition, relatively good NPBS available today tend to have either narrow useful bandwidths, or narrow acceptance angles (as small as ±1°). While this is far from ideal, the previous generation used thin metal films, and tended to be about 25% diattenuating.

Properties of polarizers are highly dependent on the particular type of polarizer. At normal incidence in a perfectly collimated beam, all polarizers will behave as expected. For non-normal use, the properties are so variable that a survey is not practical here. As a practical matter, all polarizers have an acceptance angle of at least ±1°, with some considerably larger than this.

Retardation plates, frequently called waveplates, are commonly made from thin plates of birefringent material. A waveplate which is actually the retardance specified, for example 0.25 waves, is generally too thin to mechanically support itself, and is therefore glued or grown onto a plate of glass. These minimize the variation of retardance with angle of incidence and wavelength. Some older books refer to these as zero order waveplates, though today they are called "true zero order". Plates which are thick enough to be self-supporting are called high order waveplates. A high order quarter waveplate will not be 0.25 waves of retardance, but actually some higher number, such as 10.25 or 30.25 waves, and at normal incidence in monochromatic light, the effects will be the same. The spectral and angular variation are, far more rapid than a true zero order waveplate. Another common class of waveplate is today called a "zero order waveplate", though older works may refer to it as a "first order waveplate". This is two high order waveplates, where the retardance of the two plates differs by a specified amount, glued such that the fast axis of one is aligned with the slow axis of the other. This makes the spectral variation the same as a true zero order waveplate, but the angle of incidence variation is still that of a high order waveplate.

A word about the variation of retardance with wavelength and angle of incidence is in order. Neglecting higher order effects such as dispersion, the retardance, in nm, of a particular path through a birefringent medium is constant for all wavelengths. This leads to the retardance, normalized to the current wavelength, scaling inversely proportional to wavelength $\delta = \delta_0 \dfrac{\lambda_0}{\lambda}$. The dependence of retardance on angle of incidence is rather more complicated. Suppose a ray is fixed in space, nominally passing at normal incidence through a plane-parallel block of birefringent material. The ray experiences a certain retardance, δ_0. Now rotate the birefringent plate about its fast axis. For moderate rotations, the retardance increases quadratically with angle of incidence. For moderate rotations about

the slow axis, the retardance decreases with a gentler quadratic curve. This results in an approximately toroidal retardance variation as the angle of incidence on the birefringent medium changes. As an example, a quartz plate 1.973 mm in thickness, undergoing a 10° rotation about the fast axis, results in a "quarter wave" retarder (actually 10.25 waves) approaching a "half wave" retarder (10.43 waves, half wave being 10.5 waves). Due to these properties, a waveplate with a spherical wave incident will show an aberration pattern which appears to be a combination of retardance astigmatism, retardance defocus, and retardance piston.

Design of Polarization Critical Optical Systems:

From the preceding sections we can see that in real optical systems, every single element causes polarization problems, whether diattenuation or retardance. For this reason, it is desirable to not have optics in a polarization system. For some systems this is feasible. Consider the sample measuring polarimeter shown in Figure 14. In this configuration, there is a high quality polarizer after the illumination optics, so that any polarization effects in the illumination system are unimportant, as long as they are constant with time. Similarly, there is another high quality polarizer as the last element in the polarization analyzer, putting the same polarization state onto the imaging system at all times, rendering any polarization effects in the imaging system unimportant. By this technique, this system bypasses all the problems with polarization in the ordinary optical elements.

| Illumination Optics | Polarization Generator | Sample | Polarization Analyzer | Imaging Lens and Camera |
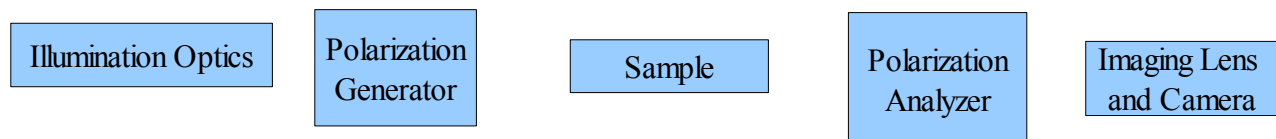
Figure 14: Block Diagram of Sample Measuring Polarimeter

For other systems, it may not be possible or practical to entirely avoid having optics in the polarization critical path. These systems are the ones for which polarization engineering becomes difficult. The comparisons with wavefront aberrations suggest applying traditional lens design approaches to solving the polarization aberrations. There is a complication, however. Whether a lens is positive or negative, if placed in a collimated beam the resulting diattenuation and retardance spherical aberration and defocus will have the SAME sign, not opposite signs as they would with traditional scalar wavefronts. This means for a rotationally symmetric optical system, all components add to the polarization errors, none subtract (excepting certain clever thin films, discussed later). For the retardance and diattenuation to subtract, somehow the s & p vectors of the surfaces must be reversed.

One way to accomplish this is through crossed fold mirrors. Crossed fold mirrors, in very low NA beams, very nicely reverse the s & p – polarizations. For symmetric optical systems with low NA sections in the middle, such as the relay in Figure 17, this is a viable technique. This technique does increase the bulk of the system, and make mounting more difficult. Also, as seen earlier for non-planar wavefronts there is a predominantly linear variation in the polarization properties of a fold mirror. When two are crossed, the diattenuation and retardance piston terms cancel, the quadratic terms add, and the linear terms couple together and create a diagonal linear term whose magnitude is given by the root-sum-square of the two (patten for crossed fold mirrors shown in Figure 16). It is helpful to use penta-prism arrangements instead of 45° fold mirrors to minimize these effects.
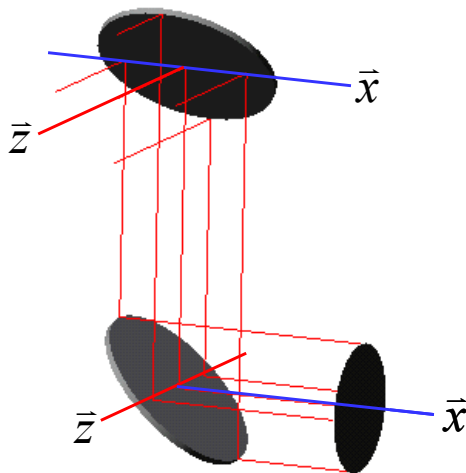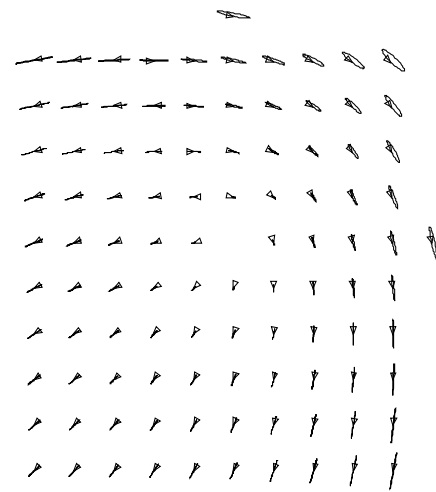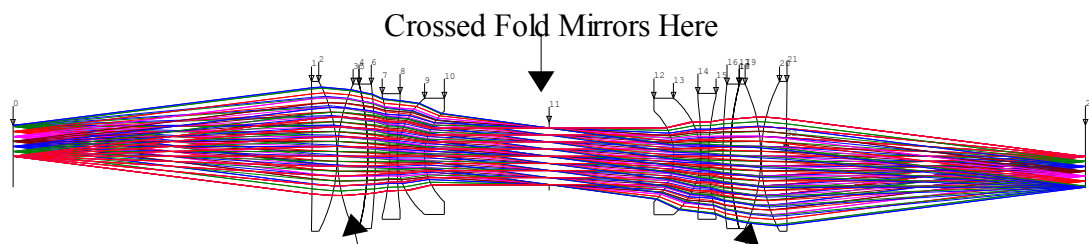

Figure 15: Crossed Fold Mirrors

Figure 16: Polarization Aberration
Pattern for Crossed Fold Mirrors

Crossed Fold Mirrors Here

Then Polarization Effects From These Two Approximately Cancel

Figure 17: Symmetric Relay Lens

Another option is applying classic low stress optical design methods discussed by Sasian and others which tends to produce low polarization optical systems. It can be shown that a classically low stress optic has small angles of incidence everywhere. This is exactly the condition for small polarization effects. Achieving low stress optical design nearly always increases the number of

elements in the system. It can certainly increase the cost and complexity of the system, but it tends to reduce the polarization aberrations since polarization aberrations are generally quadratic in angle of incidence.

There is an interesting element which is a block of uniaxial birefringent material where the extraordinary index is oriented along the axis of propagation. This makes the element optically isotropic at normal incidence, with a quadratic retardance for spherical wavefronts. Since uniaxial materials are available with both positive and negative birefringence, it is possible to pick the sign of the resulting retardance defocus by appropriate material choice.

This paper has neglected the field dependent polarization aberrations. This is due to the fact that, at present, there are zero published techniques for compensating field dependent polarization aberrations. This leaves the optical designer only the option of minimizing the polarization aberrations through whatever techniques are plausible for that particular system.

In systems designed to operate in the ultra-violet, all transmitting materials are birefringent, even fused silica. Techniques to minimize the effects of using birefringent lenses have been pioneered by the lithography industry. They rely heavily on rotating lenses about the optical axis to particular orientations. For details on this, see published literature.

Other points which occasionally may be useful for the prospective polarization system engineer. NPBS were commented upon earlier. Since these elements are so non-ideal, and can vary so rapidly, it is a good idea, whenever possible, to arrange the system assembly process so that the NPBS can be aligned, then staked in place, then measured in exactly the orientation in which it will be used. This allows reasonably accurate removal of the effects of the NPBS on any measurement by software algorithms.

What remains after the above has been completed is largely a complicated, iterative exercise in thin film design. Thin films exist which reverse the usual properties of spherical surfaces, making the s-polarization state more transmitting than the p-polarization for example. Thin films cannot change the critical angle of a prism, but the retardance of the TIR surface can be drastically altered. In theory, one could make any optical system polarization corrected, simply by creative manipulation of the thin film coatings. In practice, this turns out to be problematic. Coating designs can quickly become quite complex, and certain combinations of polarization properties simply have not yet been obtained. A good example of this is a modern NPBS. The coating may be more than 200 layers thick, and the residual polarization properties are still a significant impediment to high accuracy polarization systems.

Also, no commercial software exists which can simultaneously optimize the geometrical optical design and the coating design, in order to achieve a low polarization system. All such efforts this author has witnessed have been extremely labor intensive, utilizing copious amounts of custom software tailored to the particular problem at hand. Even a three element system can easily consume hundreds of hours of personnel time to truly make a low polarization design.

Once the design of a polarization sensitive system has been finished, it must be built. At the present time very few manufacturers and coating houses have any significant capability to measure the polarization properties of their products. As a result, received products quite often do not have the expected properties. This is true especially of thin films. Multiple iterations of the design-build-measure cycle are commonly needed to achieve really low polarization designs.

Conclusion:

This paper has provided an introduction to sources of polarization errors in optical systems. It then discusses particular optical elements and characteristic polarization aberrations experienced by those elements. Finally, the paper provides, using the above resources, a discussion of techniques for reducing the polarization properties of optical systems. Intertwined are brief discussions of what is necessary for mounting certain optical elements in polarization critical systems. Finally, on the next page is a bulleted list summarizing the major points from this paper, followed by various useful polarization references.

Bulleted Summary and Rules of Thumb:

Bulleted Summary:

- Major sources of polarization effects
  - Birefringence
  - Dichroism
  - Fresnel Effects
  - Thin Film Coatings
- Effects of Optical Elements
  - Windows
    - Can use as polarizer
    - Usually better to place normal to incident wavefront
  - Lenses & Mirrors with Power
    - quadratic and quartic polarization aberrations
    - Unless odd thin films are used, all aberrations add in a rotationally symmetric system
  - Fold Mirrors & Prisms
    - Large polarization aberrations
    - Reduce by using two mirrors in pentaprism configuration
  - Beam Splitters
    - Only use in collimated space
    - A Good NPBS doesn't exist
    - PSB are pretty good
  - Polarization Critical System Design Approach
    - Keep optics out of the polarization critical region of any system
    - Keep angles of incidence small
      - Traditional low-stress optical design techniques are helpful for this
    - Avoid NPBS if possible
      - If not, try to measure in place

Rules of Thumb:
1. Polarization properties are usually quadratically proportional to the angle of incidence
2. When in doubt about the angular acceptance of a polarization element, 1° is a safe number
3. A glass surface in TIR has $1/8^{th}$ wave of retardance

References:

1. Chipman, Russell A.   Dissertation, University of Arizona, 1987
2. Shurcliff and Ballard, "Polarized Light", Van Nostrand Co, 1964
3. Kliger & Lewis, "Polarized Light in Optics and Spectroscopy", Academic Press, 1990
4. Goldstein, "Polarized Light", Marcl Dekker, 2003
5. OSA Handbook of Optics, chapters on Polarization, Polarizers, and Polarimetry
6. McClain et. al., "Polarization ray tracing in anisotropic optically active media. II. Theory and physics", Vol. 10, No. 11/November 1993/J. Opt. Soc. Am. A
7. Jackson, "Classical Electrodynamics", 3rd Ed, Wiley, 1999
8. Macleod, "Thin Film Optical Filters", 3rd ed, IoP, 2001