

Optical Design of Multilayer Achromatic Waveplates by Simulated Annealing and their Applications

J. Ma, J. Wang, T. Spirock, C. Denker, and H. Wang

Big Bear Solar Observatory, 40386 North Shore Lane, Big Bear City, CA 92314

New Jersey Institute of Technology, Center for Solar-Terrestrial Research, 323 Martin Luther Kind Blvd., Newark, NJ 07102

jxm1901@njit.edu

ABSTRACT

In this paper we applied the simulated annealing algorithm to solve an optimization problem to design multilayer achromatic waveplates using uniaxial crystals. We derived the results specifically for the wavelength range from 1000 nm to 1800 nm, which is the wavelength range for the Near Infrared Tunable Filter system (NIRTF) of Advanced Technology Solar Telescope (ATST). The retardance of six- and ten-layer achromatic quarter waveplates within this wavelength range is found to be $89^{\circ}48'35'' \pm 0^{\circ}56'8''$ and $89^{\circ}52'18'' \pm 0^{\circ}35'28''$ respectively. We then transferred this multilayer waveplates design to visible wavelength range. The optical performance of multilayer waveplates are discussed by considering several simulated experiments. The example applications of multilayer achromatic waveplates are also discussed, specifically, for a general purpose phase shifter and the birefringent filter in the Infrared Imaging Magnetograph (IRIM) system of BBSO. In Appendix C, we also discussed in details a method to measure the retardance of waveplates.

Subject headings: Sun: infrared — instrumentation: polarimeters — methods: numerical

1. Introduction

The Sun is the only star where we can observe the fine structure of stellar magneto-convective motion, which is one of the most important tasks in understanding stellar structure and evolution. The ATST (Keil et al. 2001, 2005; Rimmele et al. 2005) is a 4-meter off-axis ground based telescope proposed for observations of solar fine structures. ATST will achieve an angular resolution of $0.03''$ (visible), and cover the spectral range from 0.35 to 35 μm with minimized scattered light and high order adaptive optics (Rimmele 2004). A set of focal plane instruments is currently being designed to exploit the high resolution capabilities of ATST (Rimmele et al. 2004). One of the proposed focal plane instruments is the Near Infrared Tunable Filter (NIRTF) system, which will be discussed briefly in the current paper.

A tunable filter system provides astronomers the unique opportunity to observe the hydrodynamic motions and magnetic fields on the Sun in 2D at different wavelengths with moderately high spectral resolution ($R = \lambda/\Delta\lambda \approx 10000 - 50000$). For example, the existing Universal Birefringent Filter (UBF) at National Solar Observatory/Sacramento Peak (NSO/SP), New Mexico has a tuning ability from 400 to 700 nm with a bandpass of 0.05 or 0.025 nm (two operation modes) (Beckers et al. 1975). The proposed NIRTF system will achieve a tuning ability across the spectral range from 1000 to 1800 nm and a combined bandpass as narrow as 4.5 pm corresponding to a spectral resolution of $R = 180000 - 250000$ at 1560 nm. By observing the Zeeman effect of the magnetic fields on the Sun at spectral lines around 1.56 μm , we are able to measure the magnetic structure and evolution in the deepest visible layer of the Sun, i.e., the lower photosphere. Due to the larger Zeeman splitting at this wavelength, we can also detect the weak fields distribution (Lin & Rimmele 1999).

Solar 2D spectro-polarimetry usually follows one of two

different approaches, i.e, either a scanning spectrograph or a narrow-band filter image system records the 3D data cube consisting of one spectral and two spatial dimensions. The choice is to scan one of the spatial dimensions (spectrograph) or the spectral dimension (imaging systems). The latter approach of obtaining the data is particularly well suited for post-facto image restoration, if the exposure times can be kept short compared to the correlation time-scale of the seeing (about 40 ms under good day-time conditions). High spectral resolution is one of the advantages of a spectrograph. However, scanning a large field-of-view (FOV) with a narrow entrance slit would take a long time. For example, a typical observation of the Advanced Stokes Polarimeter (ASP) over a FOV of $165''$ (w) by $90''$ (h) will last for half an hour (Kubo et al. 2003). Thus, dynamic features such filaments, spicules, and other fast evolving hydrodynamic plasma motions are difficult to capture using spectrograph systems. Typically, we want to be able to follow the fast changing solar features on time-scales of 1 min or less. Filter systems provide exactly these capabilities: large FOVs and fast spectral scans. A brief description of the scientific objectives of the NIRTF is listed in Table 1, which details interesting spectral lines from 1000 to 1800 nm. Take Fe I 1564.85 nm for an example, by operating NIRTF in the third mode (narrow band) listed in Table 1, one can obtain the spectral profiles of the Stokes components of Fe I 1564.85 nm. From the analysis of the Stokes profiles, one can derive field strength, filling factor for sub-gauss magnetic field (Rabin 1992). If NIRTF is operated in the second mode (i.e., with blocking filter and birefringent filter only), the magnetic flux or averaged field strength can be derived corresponding to the field strength of **sub-gauss for longitudinal field, and 10 G field for transversal field**. Compared with the observations in the visible spectrum, the seeing conditions at infrared wavelength also benefit this kind of observations.

We will use NIRTF as an example application to explain the motivation of the current work on the computational design

of multilayer achromatic waveplates. In the end, several other practical applications of such achromatic waveplates in astronomical instrumentation will also be discussed.

2. The Birefringent Filter and Spectral Tuning

The NIRTF is a group of several filters: a blocking interference filter, a three-stage tunable birefringent filter, and a tunable Fabry-Pérot interferometer (FPI). The transmission profiles of these filters in NIRTF are shown in Figure 1, in which the unique transmission property of each filter is compared. In Figure 1b, we depict the transmission profile of the birefringent filter, which has periodic transmission maxima at about every 5 nm. The interference filter is used to suppress these sidebands (see Figure 1a). If the width of the interference is properly chosen, i.e., the relatively steep flanks of the filter coincide with the first transmission minima of the birefringent filter, the first and higher order neighboring maxima of the sinc² function can be effectively suppressed. The transmission maxima of both the birefringent filter and the Fabry-Pérot interferometer (Figure 1c) are tuned in tandem from the current wavelength position to the next. In this case, the combined transmission maximum will swipe through the wavelength region containing the spectral line(s) of interest. Thus, we can assemble a spectral line profile for any point in the field-of-view (FOV). Tuning the FPI is typically performed by adjusting the distance between two parallel flat mirrors. Even though changing the refractive index in between the plates by varying the temperature or pressure has also been used in the past. This in turn means that controlling the FPI environment is essential for a stable NIRTF operation. A detailed description of the alignment and calibration procedures for FPIs was recently presented by (Denker & Tritschler 2005). In the following sections, we will describe achromatic waveplates, which play a crucial role in tuning the bandpass of the birefringent filter.

Birefringent material such as calcite has two distinct refractive indices. If a monochromatic coherent plane wave enters a birefringent material, it splits into two beams with refractive indices n_o and n_e for the ordinary and extraordinary ray, respectively. On the exit, a phase difference will be present, since the light propagates at different speeds along the ordinary and extraordinary path. This phase difference can be written as $2\pi\mu d/\lambda$, where $\mu = n_e - n_o$ defines the birefringence of the material, d is the thickness of the material and λ is the wavelength of the monochromatic wave. Calcite and quartz are among the most commonly used birefringent crystals. Note that the birefringence of calcite is much larger than that of quartz, while calcite (CaCO₃) is much softer than quartz (SiO₂) with respect to the physical hardness. Therefore, quartz can be made in very thin layers and is more suitable for the compact optical designs such as achromatic waveplates or phase shifters. These are just two of the considerations, which have to be taken into account constructing a birefringent filter made of these two materials.

A birefringent filter usually holds several stages of similar structure. Each stage is composed of blocks of calcite, linear polarizers (films or liquid crystals) and $\lambda/4$ and $\lambda/2$ waveplates (made of quartz) or liquid crystal variable retarders (LCVR). The bandpass of the filter can be shifted by mechanically rotating some of the waveplates or varying the voltage applied to the LCVR (Wang et al. 2001; Spirock et al. 2001). In the Figure 2, a schematic single stage birefringent filter is shown, which uti-

lizes a rotating waveplate to shift the bandpass. The core unit of the filter is a $\lambda/2$ waveplate sandwiched in between a split calcite of width $d/2$. This “wide-field” arrangement of the calcites (Evans 1949) allows larger FOVs and/or optical set-ups with faster beams. The phase shifting unit will be discussed in detail in the following sections. Its sole purpose is to adjust the phase difference between the two perpendicularly oscillating plane waves by a specific amount.

For a single stage of a multi-stage birefringent filter, the transmission can be represented by

$$\tau_i = \cos^2 \left(\frac{\pi\mu d_i}{\lambda} + \delta_{p_i} \right) , \quad (1)$$

where d_i is the thickness of the calcite and the subscript i refers to a specific stage of the filter. The small phase change δ_{p_i} represents the retardance adjustment by the phase shifter. In reality, δ_{p_i} could also be due to manufacturing errors in d_i or temperature changes, which in turn affect the value of μ . All these effects can be compensated by the phase shifting unit shown in Figure 2, since the phase changes produced by the phase shifting unit vary continuously from 0 to 2π . This full-period change also enables tuning the bandpass from one wavelength position to another. The bandpass shift can also be understood from the relation between δ_{p_i} and λ in Equation 1, since any change in δ_{p_i} can easily be converted to an equivalent change in λ . The same argument applies to d_i as well.

The transmission profile of a multi-stage birefringent filter is simply realized by multiplying together the τ_i of each stage. The ratio of the calcite thickness d_i in each stage has to follow the sequence 1 : 2 : 4 : 8 . . . for a perfect match to achieve the transmission profile shown in Figure 1b. Unfortunately, this perfect ratio can not be achieved in practice due to manufacturing errors. The resulting mismatch in the d_i ratios has to be compensated using phase shifters. Therefore, in addition to tuning the bandpass of the birefringent filter, the phase shifters in each stage also function as thickness compensators for the calcites. Each stage must be calibrated separately before assembling the multi-stage filter.

One of the most prominent advantages of using rotating waveplates, as opposed to LCVRs, is that the phase change δ_{p_i} produced by the rotating $\lambda/2$ waveplate is linearly related to the rotation angle of the $\lambda/2$ waveplate (see Section 5.1). In addition, the temperature stability and mechanical reliability of waveplates, which are usually made from quartz (SiO₂), are superior compared to liquid crystals. Therefore, as long as the waveplates maintain exactly the retardance relation $\lambda/4 \leftrightarrow \lambda/2 \leftrightarrow \lambda/4$, the phase shifting unit in Figure 2 is an ideal solution to tune the bandpass of a birefringent filter. In reality, a waveplate made of natural materials such as quartz can only produce a predefined retardation at a single wavelength. For example, a $\lambda/4$ waveplate designed at 1564.8 nm will not produce $\pi/2$ retardance at 1083.0 nm. The reduction of the chromatism (or wavelength dispersion) of waveplates will be the primary focus of this paper.

3. Designing Achromatic Multilayer Waveplates

There are basically two approaches to the optical design of achromatic waveplates: combinatorial and structural waveplates. A combinatorial waveplate (see Figure 3) is an optical train of slabs of birefringent materials. Both the thicknesses d_k and the azimuth θ_k are to be specified to reduce the wave-

TABLE 1
NIRTF OBSERVING MODES

Observing Modes	Bandpass	FOV	Solar Lines (nm)	Scientific Interests
Blocking Filter	20 ~ 40Å	–	None	Broad band white light imaging of active regions
Blocking Filter & Birefringent Filter	2.4 Å@ 1564.85 nm 1.18 Å@ 1083.0 nm	~ 120"	FeI-1564.85, 1565.24, HeI-1083.0, etc. CN at J-band (Asensio Ramos et al. 2005) OH at H-band	Line core/wing images
All Three Filters	0.085 Å@ 1565.0 nm, 0.045 Å@ 1083.0 nm	~ 60" (under discussion)	FeI-1572.36, 1565.29, 1564.85, 1524.50, 1520.75, 1521.96, 1287.98, 1188.41, 1188.28, 1178.33, 1160.76, HeI-1083.0, MgII-1095.2, 1091.4, etc.	Line Profiles Full Stokes Profiles Magnetograms Velocity maps from Stokes-V zero-crossing

NOTE.—See Wang et al. (2004).

length dependence of the retardance. This is the method used in this study and we will refer to them as “ n -layer” or multilayer waveplates. A structural waveplate is a single layer of material, not necessarily birefringent, but with embedded periodic surface structures on micron- or nano-scales. Such small structures can produce the so-called form birefringence (Born & Wolf 1999), which can be used to build achromatic waveplates (Kikuta et al. 1997; Flanders 1983). This approach relies on semiconductor manufacturing technology, which is beyond the scope of this presentation.

The design of an n -layer waveplates is rather complicated with respect to the computational burden. Each layer has three free parameters: thickness, azimuth angle and material property (birefringence). Therefore, the possible combinations of these parameters increase exponentially as the number of layers increases.

The analytical design of three-layer achromatic waveplates can be found in Title (1975). For a three-layer $\lambda/4$ waveplate, the author specified retardances of 115.5° , 180° and 115.5° and azimuth of 0° , 70.6° and 0° . For a three-layer half waveplate, the retardance are 180° , 180° , 180° and the azimuth are 0° , 60° , 0° . This analytical method considers the relative frequency instead of wavelength to simplify the discussion, and try to find universal configurations for three-layer achromatic waveplates regardless of the specific spectral range. A different computational approach of designing six- and ten-layer achromatic waveplates was presented by McIntyre & Harris (1968), who discussed both numerical and experimental results. In this unique design, all waveplates have the same thicknesses but different azimuth. The achromatic range considered was 400 to 800 nm. Various other designs exist, which use two different birefringent materials (for example, Hariharan 1995) or prisms (Filinski & Skettrup 1984).

3.1. Definition Of The Optimization Problem

In Figure 3, let $\mathbf{J}_{\text{awp}} = \mathbf{J}_1 \cdot \mathbf{J}_2 \cdot \mathbf{J}_3 \cdot \dots \cdot \mathbf{J}_n$ be the total Jones matrix of the multilayer waveplate and $\mathbf{J}_{\text{perfect}}$ be the Jones matrix of an ideal waveplate with a predefined retardance. Each \mathbf{J}_k depends on a set of parameters θ_k and δ_k with $k = 1, \dots, n$, which are the azimuth angle and the retardance of each layer, respectively. The design criterium of an n -layer achromatic waveplate

is to minimize the following merit function

$$E(\theta_k, \delta_k) = \sum_{\lambda, i, j} \Delta J_{ij} \cdot \Delta J_{ij}^* |_{\lambda} \quad (2)$$

$$\text{with } \delta_k = 2\pi\mu d_k / \lambda \quad (3)$$

The matrix $\Delta \mathbf{J} = \mathbf{J}_{\text{awp}} - \mathbf{J}_{\text{perfect}}$ represents the Jones matrices difference between an n -layer waveplate and an ideal waveplate, which has four elements. Note that the $\delta - \lambda$ relation assumed in the above merit function is only valid for natural crystals such as quartz and calcite. Other birefringent materials, for example the aforementioned liquid crystals, have different $\delta - \lambda$ relations. One simple way of arranging the summation over the wavelength λ is to sample a large number (> 100) of equally spaced points between 1000 to 1800 nm. This arrangement of summation implies that every wavelength within the range is equally weighted. If there are any spectral lines of particular interest, then the merit function can be evaluated only at these locations such that the final design will be particularly optimized in the vicinity of these spectral lines. In general, the merit function can be multiplied by a weighting function that reflects the specific scientific objectives of the instrument.

The merit function defined in Equation 2 has $n \times 2$ free parameters and is highly non-linear. A cursory investigation of this function shows that it possesses a very large number of quenches, i.e., local minima. Many of the standard minimization methods for global optimization problems with multiple parameters fail to locate and rank these local minima. In this study, we apply the Simulated Annealing (SA) algorithm (Kirkpatrick et al. 1983) to solve the combinatorial optimization problem defined by Equation 2 and 3.

3.2. The Simulated Annealing Algorithm

The SA algorithm is a generalized Monte Carlo numerical simulation algorithm first introduced by Kirkpatrick et al. (1983) and has been successfully applied to many scientific and engineering problems (Habib et al. 2006; Sun et al. 2005). For the details of the development of the SA algorithm, we refer to the original work by Kirkpatrick et al. (1983), the monograph regarding Monte Carlo method by Newman & Barkema (1999) and also a brief description in Appendix B. In the annealing process of melted materials, the system in equilibrium described

by Gibbs statistics approaches the ground energy state when the annealing temperature approaches zero. Thus, an analogy can be drawn between the crystal annealing process and the combinatorial optimization problem.

At the starting point of an SA program, the system represented by the energy E (parameterized by a set of free parameters) is put at a high temperature (“effective temperature”), e.g., $T_0 = 5$. Note that this temperature is only a concept borrowed from statistical mechanics, which has nothing to do with real temperature. It is in fact a measure of the magnitude of the random move in the state space. This temperature is slowly reduced based on a predefined annealing schedule. At each temperature, the system is allowed to stay long enough to achieve an equilibrium state while the state space (a nonconvex hyperspace of free parameters) is sampled according to a sampling probability. For details, we refer to the Appendix B.1. Each state in the state space corresponds to an energy state of the system. The transition from one energy state E_1 to another E_2 is controlled by the Metropolis algorithm — If $E_2 < E_1$, the move in the state space will be accepted. Otherwise, it will be accepted with the probability given by

$$P_{12} = \exp\left(-\frac{E_2 - E_1}{T}\right), \quad (4)$$

where E_2 is the energy of the next state and E_1 is the energy of the previous state. It can be shown that any Monte Carlo simulation utilizing the Metropolis algorithm will statistically converge to a certain steady state (as long as it exists), since the algorithm satisfies the condition of detailed balance.

In SA algorithm, the concept of “energy” is defined in the most general sense as a merit function, which is to be minimized. We modify the “energy of system” defined in Equation 2 by including a weighting factor matrix \mathbf{w} in the problem definition

$$E = \sum_{\lambda} \sum_{ij} w_{ij} \Delta J_{ij} \cdot \Delta J_{ij}^*, \quad (5)$$

where the weighting factors are set to be $w_{12} = w_{21} = 100$ and $w_{11} = w_{22} = 1$. The weighting factor matrix \mathbf{w} is beneficial to improve the efficiency and accuracy of the optimization due to the following reasons. The two anti-diagonal matrix elements of the Jones matrix of an ideal waveplate are precisely zeros. Therefore, for a multilayer achromatic waveplate, any non-zero residuals in these two matrix elements will introduce significant (optical) errors in the final achromatic waveplate. The weighting factors will guide the SA simulation to identify “energy states” with minimized anti-diagonal matrix elements. The design results in Figure 4 are based on this successful strategy.

As discussed by Kirkpatrick et al. (1983); Munakata & Nakamura (2001), the annealing schedule is critical to the performance of the algorithm. We took a very simple approach to schedule the annealing process. Whenever we find the system energy stalled at one temperature after 20000 function calls to the merit function, we reduce the temperature by multiplying a factor, i.e.,

$$T_{t+1} = cT_t \quad (6)$$

where, the subscription t records the occasions of decreasing the temperature. In our computing program, we chose $c = 0.95$. Although this is not a rigorous treatment, it produces reasonable results. In the design problem of a ten-layer waveplate, there are 20 free parameters. It normally took the program one million

function calls to cool down close to $T = 0$. As suggested by Kirkpatrick et al. (1983), if the energy stays the same after the temperature drops three times, the program will be terminated. At this point, we consider that the global minimization of the merit function has been reached.

A sampling probability of the Markov chain in Monte Carlo algorithm has to guarantee that every state in state space has its opportunity to be scanned. This property is called the condition of ergodicity (Newman & Barkema 1999). In the current study, we simply changed one free parameter by a random amount (either positive or negative) in one function call to the merit function at a time. This is equivalent to a uniformly random scan of the state space.

3.3. Comments On The Computing Results

In the paper by Title (1975), the author showed analytically that, using same uniaxial material, at least three layers are required in order to form a multilayer waveplate. In the SA simulation, this conclusion is confirmed since we observed that a two-layer waveplate system will not reach a sufficiently low energy ground state, even if the effective temperature has been reduced to zero. For a multilayer waveplate with more than two layers, the energy reduces constantly as the temperature decreases. The results of the SA algorithm are collected in Table 2, which includes three-, six- and ten-layer achromatic waveplates designs.

Note that in the three-layer $\lambda/4$ waveplate configuration in Table 2, the first and third layers have very similar parameters. Similarly, all layers of the three-layer $\lambda/2$ waveplate have roughly the same retardance. These results agree with the results based on the analytical method in Title (1975). However there exists quantitative differences between the computed configurations and analytical counterparts. This is due to the different definitions of the optimization problem. As defined by Equation 2 or Equation 5, the error of the computed configurations are “globally” optimized from 1000 to 1800 nm. The analytical configurations were optimized locally.

In the following section, we will discuss the optical properties of the configurations listed in Table 2.

4. Numerical Evaluation of The Optical Performance of n-Layer Achromatic Waveplates

In an optical system, which includes waveplates and polarizers, both Jones and Mueller formalisms need to be considered to fully describe the optical properties of the system. A full discussion of the development and applications of these formalisms used in polarization optics is prescribed in the monograph by Shurcliff (1962). Briefly speaking, the optical components, such as a polarizer or a waveplate, can be described by either 2×2 Jones matrices or 4×4 Mueller matrices. Correspondingly, the light is in the form of Jones vectors or Stokes vectors.

4.1. Normal Incidence

In Equation 3, the geometric length of the light path and the thickness of the crystal plate are the same, i.e., d_k , given that the light incidents on the surface of the crystal perpendicularly. In the case of normal incidence, we will carry out several computational experiments using Jones and Mueller algebra by considering an n -layer achromatic $\lambda/4$ waveplate as an example.

TABLE 2
CONFIGURATION OF THE n -LAYER ACHROMATIC WAVEPLATES DETERMINED BY SIMULATED ANNEALING

Type	#	θ_i	δ_i	θ_i	δ_i	θ_i	δ_i
$\lambda/4$ waveplates	1	149° 18' 40"	95° 58' 8"	61° 24' 39"	116° 59' 17"	24° 14' 44"	180° 11' 42"
	2	39° 18' 47"	149° 24' 7"	168° 7' 59"	150° 18' 43"	113° 3' 14"	132° 24' 54"
	3	149° 19' 19"	95° 56' 32"	100° 58' 52"	5° 18' 58"	149° 49' 37"	169° 53' 28"
	4			116° 48' 18"	198° 23' 20"	59° 39' 50"	245° 6' 47"
	5			25° 52' 26"	60° 4' 13"	66° 58' 4"	75° 11' 7"
	6			16° 33' 52"	200° 6' 0"	161° 22' 5"	183° 5' 46"
	7					131° 1' 44"	215° 13' 41"
	8					37° 2' 2"	134° 8' 20"
	9					13° 52' 33"	27° 15' 33"
	10					22° 31' 33"	73° 29' 19"
$\lambda/2$ waveplates	1	29° 24' 44"	149° 27' 4"	22° 15' 58"	138° 37' 12"	26° 9' 3"	111° 9' 47"
	2	150° 33' 32"	149° 28' 16"	58° 40' 53"	137° 43' 12"	137° 28' 55"	155° 21' 18"
	3	29° 22' 49"	149° 26' 31"	157° 39' 22"	359° 45' 50"	37° 35' 36"	239° 18' 36"
	4			39° 26' 6"	50° 13' 49"	146° 6' 58"	309° 20' 56"
	5			88° 35' 56"	65° 29' 57"	149° 6' 11"	1° 50' 21"
	6			31° 16' 37"	172° 1' 23"	49° 7' 6"	142° 25' 5"
	7					19° 33' 30"	202° 47' 28"
	8					153° 11' 20"	42° 59' 44"
	9					82° 34' 35"	34° 50' 2"
	10					102° 3' 29"	78° 53' 17"

NOTE.—All retardance δ_i shall be measured at $\lambda_0 = 1523.1$ nm. For a single waveplate, the $\delta - \lambda$ relation is assumed to be $\delta \propto \lambda^{-1}$, which is true for natural crystals such as quartz and calcite.

In the first experiment, waveplates are set up between two linear polarizers, see the experimental arrangement in Appendix C. The active angle of the first polarizer is set to 45° with respect to the optical axes of the waveplates. The second polarizer is rotating along its azimuth axis. Since the waveplates are assumed to be $\lambda/4$ waveplates, the output intensity shall be identical as the second polarizer is rotating, if the waveplates are perfect $\lambda/4$ waveplates. The computational results shows that the ratio between the maximum intensity and the minimum intensities, i.e., the ellipticity of the ellipse, is not necessarily unity — the output light after passing the waveplate is elliptically polarized (see Figure 4). Within these designs of $\lambda/4$ -waveplate, the ten-layer achromatic $\lambda/4$ waveplate has closest ellipticity to unity across the entire wavelength region from 1000 to 1800 nm. The curve has three valleys where the ellipticity is unity, i.e., perfect $\lambda/4$ waveplate. Therefore, the globally optimized solution of the Equation 5 will always have several local minima.

On the other hand, the larger diversion from unity of the ellipticity of the testing light beam shown for the three-layer waveplate is expected from a previous study (Ma et al. 2004), since the effective optical axis of a three-layer achromatic waveplate rotates to different angles at different wavelengths. A larger retardance error is inevitable. In comparison, the effective optical axes of six- and ten-layer achromatic quarter waveplates are much more stable. However, the performance of the three-layer achromatic waveplate at the two wavelength He I 1083 nm and Fe I 1565 nm is better than any globally optimized solutions for six- and ten-layer cases. The application of such three-layer waveplate in IRIM of BBSO will be investigated in Section 5.2, inasmuch as IRIM is specifically designed to carry out solar observations at the two wavelengths.

Given the ellipticity of the testing light obtained in the above mentioned method, the retardance of the waveplate can be derived by using the method shown in Appendix C. We plot the corresponding results in Figure 5, in which the retardance of single-, three-, six-, and ten-layer quarter waveplates are plotted. Across the wavelength range from 1000 nm to 1800 nm,

the maximum error of a ten-layer achromatic $\lambda/4$ waveplate is approximately 0.4% wave, or equivalently, 1.4° . As will be discussed in later sections, this error on retardance is negligible in the application of a birefringent filter, which is our main goal.

The second experiment is concerned with the evaluation of the characteristics of polarization transformation in polarimetry analysis. Linearly polarized monochromatic light is sent into the optical train as shown in Figure 3. The output light is then recorded in the form of the Stokes vector. This process has been applied to: (a) an ideal $\lambda/4$ waveplate, (b) a single layer $\lambda/4$ waveplate and (c) an n -layer achromatic $\lambda/4$ waveplates. It can be shown that the Stokes vector of linear polarized light along a direction α with respect to the x -axis is

$$\mathbf{S} = [1, \cos^2 \alpha - \sin^2 \alpha, 2 \sin \alpha \cos \alpha, 0]^T \quad (7)$$

Let \mathbf{S}_{in} be the Stokes vector of the incoming light. Then, the output is $\mathbf{S}_{\text{out}} = \mathbf{M}_{\text{awp}} \cdot \mathbf{S}_{\text{in}}$, where \mathbf{M}_{awp} is the Mueller matrix of the n -layer achromatic waveplate (see Figure 6). All the waveplates in (a), (b) and (c) are rotated by an arbitrary angle of 108° with respect to the x -axis, which simulates a practical application of rotating waveplates. The simulation shows that, after passing through $\lambda/4$ waveplates, there is re-distribution of the polarization energy among Q, U, V -components and will cause varying polarization errors, for instance, in the measurement of the vector magnetic field of the Sun (Kuhn et al. 1994). In measuring solar magnetic fields the Q - and U -components are directly related to the transverse component of the magnetic field. The V -component is related to the longitudinal component. Therefore, the crosstalk of V -component due to the polarization error of waveplates can appear as spurious magnetic signal in Q -component. In the DVMG calibration, the Stokes- V signals can be identified in the Stokes- Q image, when the optical axis and retardance of the ferroelectric crystal (FLC) is not accurate (Spirock 2005). Figure 7 is an example of the crosstalk during the calibration of the DVMG system. The spurious magnetic features in Stokes- Q image in the right panel are highlighted with contours come from Stokes- V image. Since this calibra-

tion was performed at the center of the solar disk, the Stokes-Q and -U components should be undetectable.

In the situation shown in Figure 6, the crosstalk due to n -layer waveplate is roughly limited to 2% across the desired wavelength range. This is an incomplete measure of the polarimetric accuracy of the waveplates. Note that only the linearly polarized light has been considered in this numerical experiment. In order to fully understand the polarimetric characteristics of n -layer waveplates, the elements of their Muller matrices are investigated for distinct wavelengths and for different azimuth. We will only consider the elements $\mathbf{M}(i,j)$ with $i, j = 2, 3, 4$, which define the crosstalk between linear polarizations and circular polarizations. In Figure 8, these matrix elements of a rotating ten-layer achromatic quarter waveplate is investigated at $\lambda_0 = 1083.0$ nm. And let

$$\Delta\mathbf{M} = \mathbf{M}_{10} - \mathbf{M}_{\text{ideal}} \quad (8)$$

be the difference between the Mueller matrices of the ten-layer quarter waveplate (\mathbf{M}_{10}) and an ideal quarter waveplate ($\mathbf{M}_{\text{ideal}}$). Two features can be read from this figure: (a) the magnitudes of all the elements of $\Delta\mathbf{M}$ are oscillating with respect to the azimuth of the rotating waveplate, except $\Delta\mathbf{M}(4,4)$, which is negligibly small; and (b) the period of the oscillations are different for some of the elements. The plot shows how far the performance of a ten-layer achromatic quarter waveplate is different from an ideal quarter waveplate.

4.2. Performance Stability of n -Layer Waveplates

In many circumstances, such as, dimensional inaccuracy during polishing process, or the extension/contraction due to temperature variations, the thickness of each layer of the n -layer waveplates might experience fractional changes, which will effect the retardance that each layer can provide. To simplify the discussion, consider the amplitudes of the random error (noise) are $A_\delta = 30''$ for retardance δ_k and $A_\phi = 30'$ for the azimuth ϕ_k of each layer in a n -layer $\lambda/4$ waveplate. The noise is generated from pseudo-random numbers and shifted to the range from -1 to 1. In Figure 9, the approach used in the first experiment in the Section 4.1 is performed at a single wavelength $\lambda = 1523.1$ nm for both three- and ten-layer $\lambda/4$ waveplates. It is found that the noise amplitude A_ϕ contribute a very large portion to the error of ellipticity (the retardance). The assumed noise amplitude (i.e., a measure of the accuracy) for $A_\phi = 10'$ is basically unacceptable. We found a reasonable requirement on A_ϕ should be lower than $1'$ in order to obtain a fair accuracy of the retardance of the n -layer $\lambda/4$ waveplate. This disadvantage is due to the relatively large number of layers. On the other hand, $A_\delta \leq 30''$ is sufficiently small and will not compromise the retardance of the waveplate, and is also within the controllable polishing error tolerance.

This discussion suggests when we carry out a real experiment on these n -layer waveplates, the azimuth of each layer has to be aligned carefully according to the design.

4.3. Orientation of The optical Axes of n -Layer Waveplates

The optical axes of Pancharatnam-type achromatic waveplates rotate to distinct angles with the changing wavelength (Beckers et al. 1975). This issue is minimized in the six- and ten-layer achromatic waveplate. In the criteria of the optimization problem defined by Equation 5, $\Delta\mathbf{J}$ is the difference between

n -layer waveplate and an ideal waveplate. Therefore, the minimization of ΔE corresponds to finding an n -layer waveplate, whose Jones matrix is closest to the Jones matrix of an ideal waveplate. The optical axes of the n -layer waveplate shall overlap with the optical axes of an ideal waveplate across the whole wavelength range defined in the optimization problem.

4.4. Off-Axis Effects

Let α be the incident angle, which describes the off-axis situation. For thin crystal slab, we can neglect the refraction effects on the interfaces. We also adopt the convention that the principle axis overlapping with its crystallographic axis is defined as the optical axis of the uniaxial crystal. The effective retardance of an uniaxial slab under off-axis incident light is then written as (Evans 1949)

$$\delta = \delta_0 \left[1 + \frac{\alpha^2}{2n_o} \left(\frac{\cos^2 \beta}{n_e} - \frac{\sin^2 \beta}{n_o} \right) \right], \quad (9)$$

where δ_0 is the retardance of the birefringent crystal for normal incidence ($\alpha = 0$) and β is the angle between the projection of the incident ray on the interface and the principle axis of the crystal with smaller refractive index. In literatures, this axis is sometimes referred to as the ‘‘fast axis’’, since the phase velocity along this axis is larger than the other. The fast axis is not necessarily the optical axis of the uniaxial crystal. For calcite ($n_o > n_e$), the fast axis and optical axis are perpendicular and β in the Equation 9 has to be replaced by $\beta + \pi/2$ (for a detailed discussion see Evans 1949).

Consider using quartz as the material of making the n -layer waveplates. The refractive indices of crystal quartz at different wavelength are found in Ghosh (1999). Given $\alpha = 1^\circ$ and let azimuth β vary from 0 to 360° , the maximum of the possible change on the retardance listed in Table 2 is approximately $1' 30''$ according to Equation 9. In Section 4.2, we discussed the random retardance error of each layer of n -layer waveplates due to polishing processes and environment changes. The inclined incidence discussed here can be another error source for the retardance of each layer. The error magnitude of $1' 30''$ alone will not be a serious problem for each single layer. However, if the random error and the error due to incident light are combined together, the accuracy of the retardance of n -layer waveplates could be unpredictably compromised.

4.5. Conjugate Achromatic Region

Until now, we were mainly concerned with the spectral region from 1000 to 1800 nm. The results were summarized in Table 2, where the retardance of each layer is given with respect to the reference wavelength $\lambda_0 = 1523.1$ nm. However, the parameters in Table 2 can be easily transferred to another spectral region.

If a new $\lambda'_0 = 532$ nm is specified, it can be shown that the parameters in Table 2 are still accurate in spectral region from 350 to 629 nm. We can name this region the ‘‘conjugate achromatic region’’. Obviously, by adjusting the reference wavelength λ_0 , i.e., scaling the thicknesses of all layers by the same amount, the conjugate achromatic region can be shifted. The conjugate achromatic region is scaled by a factor of λ'_0/λ_0 compared to the original achromatic range. The concept of ‘‘relative frequency’’ (Title 1975) is another approach to describe the conjugate achromatic region. From the perspective of efficiency,

a designer should design multilayer achromatic waveplates at a shorter wavelength. If the results are applied to a longer wavelengths, they will cover a much broader achromatic wavelength range. However, one should also expect that a broadband achromatic waveplate design at shorter wavelengths is more difficult using crystalline birefringent crystals.

5. Case Studies

In this section, we want to apply the achromatic waveplates discussed above to various applications in polarimetry. Some of these examples are general enough to be used in the scientific domains beyond solar physics.

5.1. Phase Shifting Unit

Phase shifting unit is an instrument that can introduce a continuous fractional phase difference between the two perpendicular directions of polarized light, see Figure 2 for the one of the possible structures of a phase shifter. Imagine there exists a birefringent crystal, which can be expanded or contracted along the x -axis by any amount. We will call such crystal a “magic birefringent crystal”. Since for a birefringent crystal, the retardance $\delta \propto d/\lambda$ (where d is the thickness), this magic birefringent crystal is capable of continuously producing any phase change by expansion or contraction. This is an ideal phase shifter, which has been discussed by Evans (1949). The author discussed several theoretical approaches to realize the phase shifter. One of these approaches consists of a combination of one $\lambda/2$ waveplate sandwiched by two $\lambda/4$ waveplates, which is especially interesting due to its flexibility. We will discuss this approach by substituting the magic birefringent crystal with a phase shifter composed of three six-layer achromatic waveplates arranged in a fashion of ($\lambda/4 \leftrightarrow \lambda/2 \leftrightarrow \lambda/4$) with properly aligned optical axes.

The optical axes of the first waveplate ($\lambda/4$) and the third waveplate ($\lambda/4$) are 45° with respect to the polarization direction of the incident light, without loss of generality, we chose it to be the x -axis. The second waveplate ($\lambda/2$) is set on a rotary stage driven by a motor and its azimuth angle θ is measured with respect to the x -axis. In the Figure 10, the Jones matrix of the phase shifter is examined. The retardance of the phase shifter is derived from the two diagonal elements of the Jones matrix, which is shown in Figure 10a. However, the diagonal elements alone are not sufficient to define the instrumental Jones matrix. The anti-diagonal elements have to be considered. For the “magic birefringent crystal”, these two anti-diagonal elements are precisely zeros. In Figure 10b, one of the anti-diagonal elements are plotted, and the solid curve represents the phase shifter made of six-layer achromatic waveplates, which produces a Jones matrix with two almost zero anti-diagonal elements. Therefore, six-layer achromatic waveplates suffice the duty in such phase shifter. As a reference, we showed a phase shifter made of three single-crystal waveplates (dashed curves). The corresponding Jones matrix has comparatively large anti-diagonal elements. Large anti-diagonal elements of the Jones matrix usually implies that the optical axis of the retarder is rotated by an undetermined angle.

Figure 10a also shows that the retardance of the phase shifter made of achromatic waveplates is linearly related to the rotation angle of the middle waveplate ($\lambda/2$). This relation has been theoretically derived by Evans (1949, see Equation VI.10).

The dashed curve shows the case of a phase shifter consisting of single-layer waveplates designed at 1523.1 nm, which is not linear at all. Although, it can be shown that a linear relation exists, if the wavelength of the incident light is exactly 1523.1 nm, i.e., the retardance of all three single-crystal waveplates are exactly $\pi/2$, π and $\pi/2$ respectively.

In some applications, the phase shifting rate is also of concern. The Newport RGV100 Series Motorized Rotation Stage has rotation speed of 720 °/s . This rotation speed results in phase shifting rate of 8 waves/s, i.e., 125 ms/wave. Phase shifters can also be liquid crystal, i.e., Liquid Crystal Variable Retarders (LCVR), switch the alignment direction of nematic molecules. Meadowlark LCVR response time to a half wave voltage is 10 ms according to the commercial documentation. The drawbacks of LCVRs is that the surface LC molecules switch faster than the molecules in the center of the LC cell depending on the thickness of a specific LC cell (usually a thin LC cell is sandwiched by two pieces of glass), and the phase change is not linearly controllable by applying voltage. The advantage of using a achromatic-waveplate based phase shifter other than a LCVR is that the former is a linear phase shifter, and more stable to hazardous environment, e.g., UV radiation (Ye 2004).

5.2. Waveplates For IRIM

Another example that the broadband achromatic waveplates can be useful is the Near Infrared Imaging Magnetograph (IRIM) at Big Bear Solar Observatory (Denker et al. 2003; Cao et al. 2005; Denker et al. 2003). The IRIM system utilizes a Fabry-Perot interferometer, two birefringent filters, and a set of wide band interference filters. The two birefringent filters are designed specifically at He I 1083 nm and Fe I 1565 nm respectively due to the chromatism of certain optical components, e.g., waveplates, polarizers and coatings on windows. The achromatic requirements of IRIM is challenging due to the two spectral lines are far away apart. In order to solve this problem for IRIM, we need to redefine the merit function as

$$E(\phi_k, \theta_k) = \sum_{\Delta\lambda_1} \sum_{ij} w_{ij} \Delta J_{ij} \cdot \Delta J_{ij}^* + \sum_{\Delta\lambda_2} \sum_{ij} w_{ij} \Delta J_{ij} \cdot \Delta J_{ij}^* \quad , \quad (10)$$

where, $\Delta\lambda_1 = 1083 \pm 2$ nm and $\Delta\lambda_2 = 1560 \pm 2$ nm define two achromatic regions, $k = 3, 6, 10$ or any other desirable number of layers. As we have seen in previous sections, the three-layer achromatic waveplates usually possess two local minima. This characteristics can be tested here. Moreover, using the same crystal material, three is the minimum number of layers that is required to form a waveplate, and it will be a timesaving task to polish only three slices of crystal material. Then we run the simulated annealing program for this newly defined merit function, and we found the following solution set:

$$\begin{aligned} \theta_1 &= 30^\circ 36' 46'' & \delta_1 &= 97^\circ 13' 31'' \\ \theta_2 &= 140^\circ 43' 41'' & \delta_2 &= 151^\circ 14' 31'' \\ \theta_3 &= 30^\circ 36' 32'' & \delta_3 &= 97^\circ 13' 3'' \end{aligned}$$

We carried out a numerical simulation for the third stage of the Lyot filter used in IRIM, and with same design parameters of the filter as being calculated in Wang et al. (2001). In this simulation, we tried the substitution of the normal single-layer waveplates (half and quarter) by both three-layer waveplates designed above and six-layer waveplates in Table 2. The transmission profiles were calculated specifically in the vicinities of 1083 nm

and 1564 nm. And the rotating half-waveplate is adopted as the bandpass shifting mechanism, in which two azimuths of the half-waveplate are calculated, 0 and $\frac{2}{7}\pi$, see Figure 11.

When n -layer achromatic waveplates are used in a Lyot filter, compared with using traditional single crystal waveplates, there are two improvements — the bandpass symmetry and the off-band scattered light level. The following can be read in Figure 11:

- (a) When single-crystal waveplates are used, the bandpass profile are symmetric only at the wavelength where the waveplates are designed, in this case, 1083 nm. Far away from this wavelength, the bandpass profile turns to be asymmetric, but still possesses certain periodicity. For instance, in the right panel of the second row in Figure 11, the bandpass is shown with one period.
- (b) The symmetry has been improved on the third row which is the case of using three-layer achromatic waveplates. However, it is obvious the scattered light level at both wavelength ranges are higher even than single-crystal case. We investigated the periodicity of the symmetry of the bandpass, and found that the symmetry is broken again between 1083 and 1565 nm, although compared with the single-crystal case, the symmetry is only twisted slightly. It is still not clear to us why the three-layer waveplates perform so badly in Lyot filter on the off-band light level. The Jones matrices errors of three-layer waveplates were minimized at 1083 nm and 1564 nm by simulated annealing procedure. We expected to see the bandpass shape in the vicinities of these two wavelengths should be close to ideal achromatic waveplates. On the contrary, we see the scattered light level is still very high.
- (c) In the case of using six-layer achromatic waveplates, both the symmetry and the scattering light have been improved. The bandpass profiles are close to the case of using ideal achromatic waveplates.

From this simulation, we can draw our conclusions regarding applying n -layer achromatic waveplates in Lyot filter. The three-layer waveplates can not be used in Lyot filter due to very high off-band scattered light level. The off-band light is out of the control of the filter design and therefore, will impair the performance of Lyot filter and the IRIM system in an unpredictable way. Six-layer achromatic waveplates can be a perfect choice for the application in Lyot filter. We did not try to use lower number layers than six, but the simulated annealing procedure is general enough to handle any number of layers for the design of n -layer achromatic waveplates.

After the substitution of the waveplates in Lyot filter with six-layer achromatic waveplates and the elimination of the chromatism on the other optical components, the working range of the 1083 nm-Lyot filter at BBSO can be expanded to the whole wavelength range from 1000 nm to 1800 nm, although currently we are only interested in using the filter at 1083 nm and 1.56 μm . With this solution for waveplates, the filter design can be greatly simplified and the filter system is compact.

6. Conclusion

In this study, we have developed a procedure based on the SA algorithm to design combinatorial multilayer achromatic wave-

plates. Using this procedure, we found that the parameters given in Table 2 can produce well performing $\lambda/4$ and $\lambda/2$ waveplates for the visible and near infrared wavelength regions. In particular, we confirmed their suitability for the NIRTF operating in the wavelength range from 1000 to 1800 nm. The performance of such n -layer achromatic waveplates was numerically evaluated by considering the residual errors in Jones/Mueller matrices and polarization errors in Stokes vectors. We also considered briefly the impact of off-axis light. By using the six- or ten-layer achromatic waveplates, an achromatic phase shifter can be constructed based on a simple structure discussed in Evans (1949). This phase shifter can continuously replicate any phase change from 0 to 2π . Therefore, a reliable tuning solution can be found for many instruments, specifically, in our case, the tunable birefringent filters. We are currently carrying out experiments on the n -layer achromatic waveplates and building the phase shifting units.

It might turn out that thin materials with stress-induced birefringence are a better choice than quartz or calcite. Wafers of stress-induced birefringent materials can be traditional polymers, such as poly-methyl-methacrylate (PMMA), polyvinyl chloride (PVC) or polycarbonate (PC) (Delplancke et al. 1995). These polymer materials have different transmission curves in different spectral ranges. Therefore, the transmission of such materials might not be as high as uniaxial crystals. Moreover, we have to be cautious about using these non-crystalline materials. The reason is that the parameters shown in Table 2 are designed based on the assumption that the “retardance vs. wavelength” relation is $\delta \propto \lambda^{-1}$. This relation might not be true for non-crystalline materials. In this case, we have to know the exact $\delta - \lambda$ relation. However, the SA algorithm is sufficiently flexible to allow the implementation of different $\delta - \lambda$ relationships. The experimental results of our achromatic waveplates are deferred to a future paper.

JM wants to thank Dr. D. G. Wang of National Solar Observatory of China (Huairou, Beijing) for the helpful discussion on the choices of the material for making the multilayer waveplates. This work was supported by NSF under grants ATM 03-42560, ATM 02-36945, IIS ITR 03-24816 and AST MRI 00-79482 and by NASA under grant NAG 5-12782.

7. Bibliography

REFERENCES

- Asensio Ramos, A., Trujillo Bueno, J., & Collados, M. 2005, *ApJ*, 623, L57
- Beckers, J. M., Dickson, L., & Joyce, R. S. 1975, *Appl. Opt.*, 14, 2061
- Born, M., & Wolf, E. 1999, *Principles of optics : electromagnetic theory of propagation, interference and diffraction of light / Max Born and Emil Wolf ; with contributions by A.B. Bhatia ... [et al.]*. Cambridge [England] ; New York : Cambridge University Press, 1999
- Cao, W., Hartkorn, K., Ma, J., Wang, J., Xu, Y., Spirock, T., Denker, C., & Wang, H. 2005, *AGU Spring Meeting Abstracts*, A8
- Delplancke, F., Bernaerd, R., Ebbeni, J., & Sendrowicz, H. 1995, *Appl. Opt.*, 34, 2921

- Denker, C., Didkovsky, L., Ma, J., Shumko, S., Varsik, J., Wang, J., Wang, H., & Goode, P. R. 2003, *Astronomische Nachrichten*, 324, 332
- Denker, C. J., Ma, J., Wang, J., Didkovsky, L. V., Varsik, J. R., Wang, H., & Goode, P. R. 2003, *Proc. SPIE*, 4853, 223
- Denker, C., & Tritschler, A. 2005, *PASP*, 117, 1435
- Evans, J. W. 1949, *Journal of the Optical Society of America* (1917-1983), 39, 229
- Filinski, I., & Skettrup, T. 1984, *Appl. Opt.*, 23, 2747
- Flanders, D. C. 1983, *Applied Physics Letters*, 42, 492
- Ghosh, G. 1999, *Optics Communications*, 163, 95
- Habib, A., Vernin, J., Benkhaldoun, Z., & Lanteri, H. 2006, *MNRAS*, 367
- Hariharan, P. 1995, *Measurement Science and Technology*, 6, 1078
- Ingber, L. 1996, *Control and Cybernetics*, 33, 25
- Keil, S. L., Rimmele, T. R., Keller, C. U., & The Atst Team 2001, *ASP Conf. Ser. 236: Advanced Solar Polarimetry – Theory, Observation, and Instrumentation*, 236, 597
- Keil, S. L., Rimmele, T. R., Oschmann, J., Hubbard, R., Warner, M., Price, R., Dalrymple, N., & The Atst Team 2005, *IAU Symposium*, 223, 581
- Kikuta, H., Ohira, Y., & Iwata, K. 1997, *Appl. Opt.*, 36, 1566
- Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. 1983, *Science*, 220, 671
- Kubo, M., Shimizu, T., & Lites, B. W. 2003, *ApJ*, 595, 465
- Kuhn, J. R., Balasubramaniam, K. S., Kopp, G., Penn, M. J., Dombard, A. J., & Lin, H. 1994, *Sol. Phys.*, 153, 143
- Lin, H., & Rimmele, T. 1999, *ApJ*, 514, 448
- Ma, J., Wang, J., Cao, W., Denker, C., & Wang, H. 2004, *Proc. SPIE*, 5523, 139
- McIntyre, C. M., & Harris, S. E. 1968, *Journal of the Optical Society of America* (1917-1983), 58, 1575
- Munakata, T., & Nakamura, Y. 2001, *Phys. Rev. E*, 64, 046127
- Newman, M. E. J., & Barkema, G. T. 1999, *Mechanics of Composite Materials & Structures*,
- Rabin, D. 1992, *ApJ*, 391, 832
- Rimmele, T. R., et al. 2005, *Proc. SPIE*, 5901, 41
- Rimmele, T. R., et al. 2004, *Proc. SPIE*, 5492, 944
- Rimmele, T. R. 2004, *Proc. SPIE*, 5490, 34
- Shurcliff, W. A. 1962, *JRASC*, 56, 269
- Spirock, T. J. 2005, Ph.D. Thesis, Big Bear Solar Observatory & New Jersey Institute of Technology
- Spirock, T., et al. 2001, *ASP Conf. Ser. 236: Advanced Solar Polarimetry – Theory, Observation, and Instrumentation*, 236, 65
- Sun, P., Yin, Y., Li, B., Chen, T., Jin, Q., Ding, D., & Shi, A.-C. 2005, *Phys. Rev. E*, 72, 061408
- Title, A. M. 1975, *Appl. Opt.*, 14, 229
- Wang, J. et al, *Critical Design of Near Infrared Tunable Filter (NIRTF) of ATST*, 2004, NJIT & NSO, <http://atst.nso.edu>
- Wang, J., Wang, H., Goode, P. R., Spirock, T. J., Lee, C.-Y., Ravindra, N. M., Ma, J., & Denker, C. 2001, *Optical Engineering*, 40, 1016
- Ye, C. 2004, *Appl. Opt.*, 43, 4007

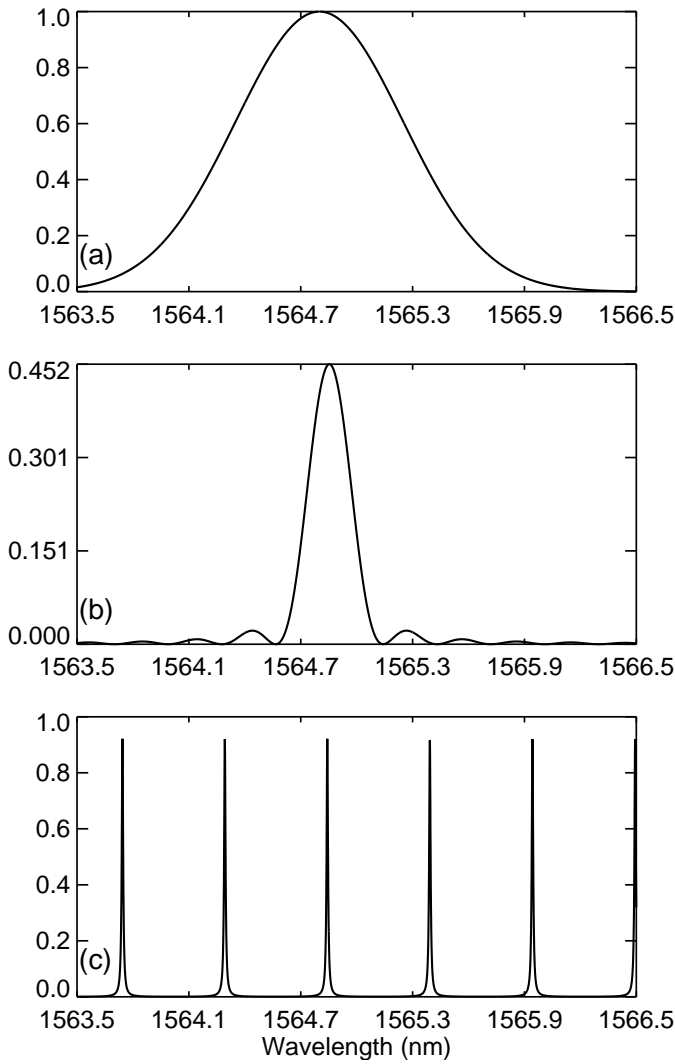


Fig. 1.— Simulated transmission profiles of the three filters restricting the bandpass of the NIRTF. (a) The blocking interference filter, (b) the tunable four-stage birefringent filter, (c) the tunable Fabry-Pérot interferometer.

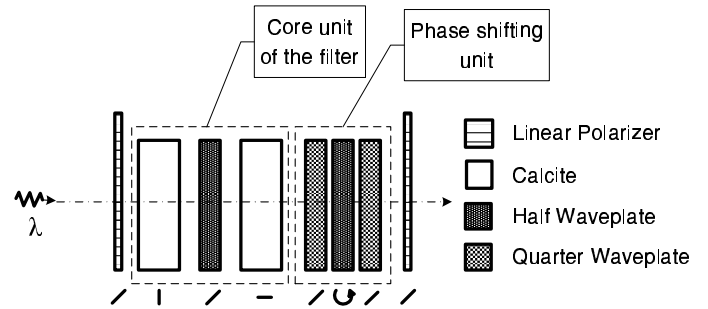


Fig. 2.— Concept design of a single stage birefringent filter with a tuning unit. The bandpass of the filter can be tuned by rotating the $\lambda/2$ waveplate of the phase shifting unit. In principle, the $\lambda/4$ waveplate next to the exit polarizer could be omitted in this filter design. It is shown here to illustrate that the phase shifting unit can be treated as an independent optical instrument in other general applications. The symbols at the bottom indicate the orientation of the (fixed and variable) optical axes of the optical elements.

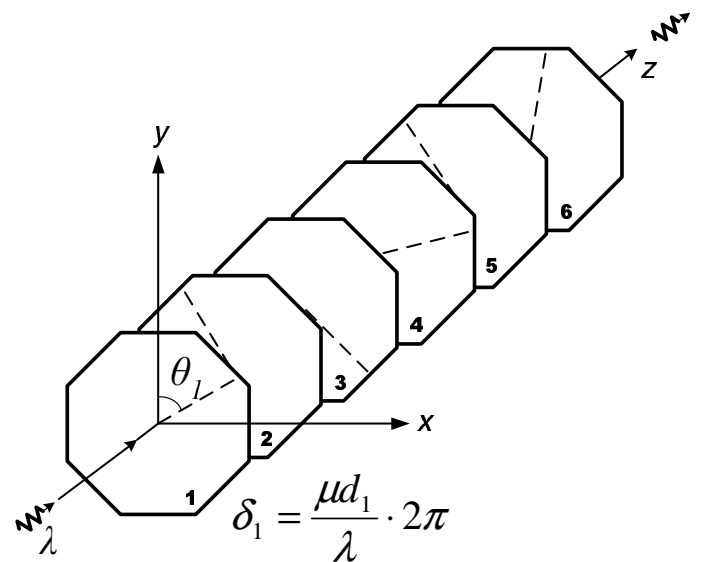


Fig. 3.— Diagram of a six-layer waveplate. Dashed lines indicate the optical axes of crystals, and are for demonstration purpose only.

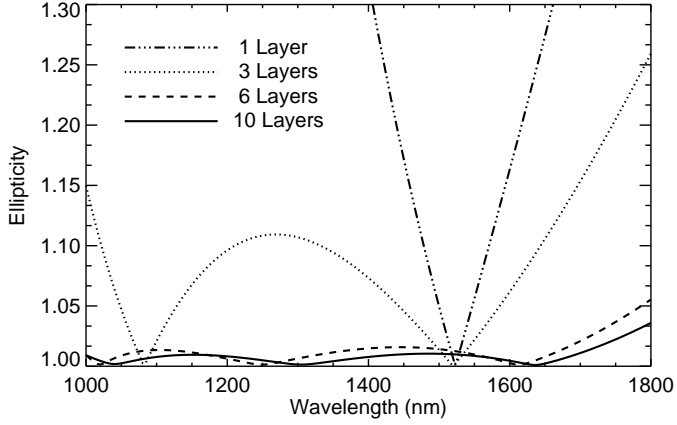


Fig. 4.— Ellipticity vs. wavelength for single crystal, three-layer, six-layer, and ten-layer $\lambda/4$ waveplates. See text in Section 4.1 for the details about the approach obtaining these curves.

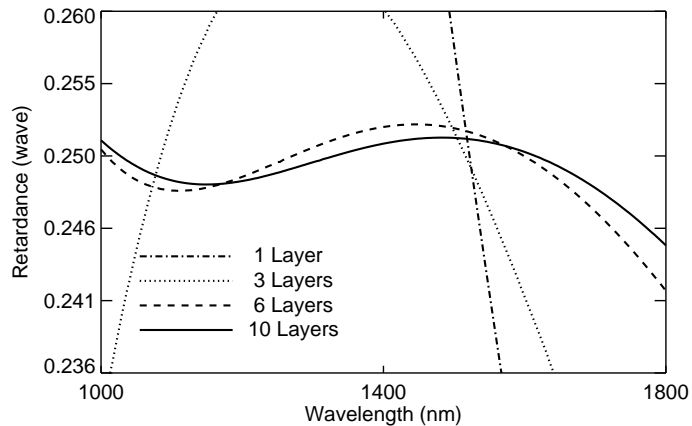


Fig. 5.— Retardance vs. wavelength relation, see Figure 4.

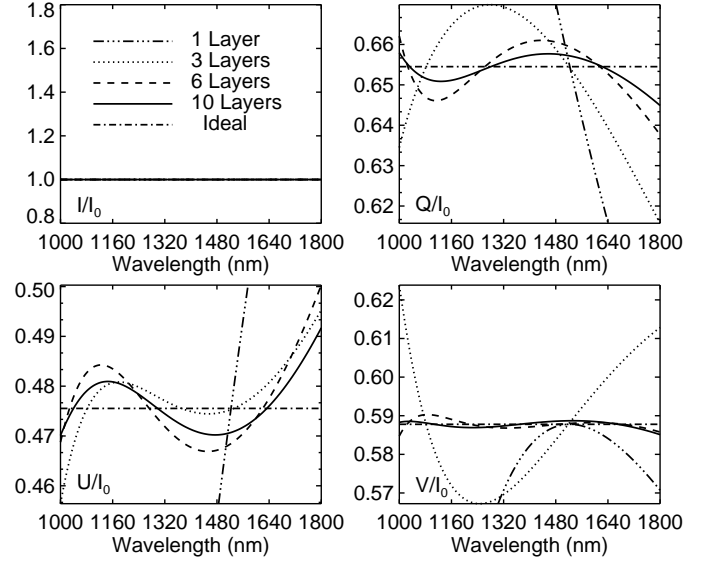


Fig. 6.— Simulated polarization crosstalk of n -layer $\lambda/4$ waveplates (see Table 2 for parameters of each layer). The effective optical axis of the overall stack of the achromatic waveplates are rotated by 108° (an arbitrarily chosen angle) to the x -axis. The incident light is linearly polarized and the angle between its polarization axis and the x -axis is equal to $\pi/5$. The corresponding Stokes vector of the light is $[1, \cos^2 \frac{\pi}{5} - \sin^2 \frac{\pi}{5}, 2 \cos \frac{\pi}{5} \sin \frac{\pi}{5}, 0]$.

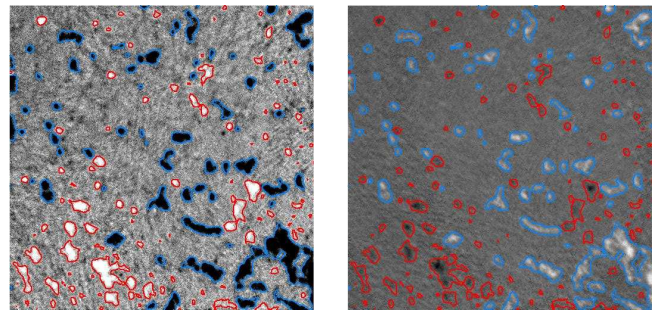


Fig. 7.— Demonstration of the crosstalk of Stokes-V and -Q (solar disk center at Ca I 630.1 nm) during the calibration of Digital Video Magnetograph (DVMG) at BBSO due to the inaccuracy of retardance and optical axes of the liquid crystal variable retarder and ferroelectric component (Spirock 2005). Left panel: Stokes-V signal; Right panel: Stokes-Q signal. Contours are used to highlight the corresponding crosstalk signals. (Courtesy of Dr. T. J. Spirock, Big Bear Solar Observatory. Reproduced with permission.)

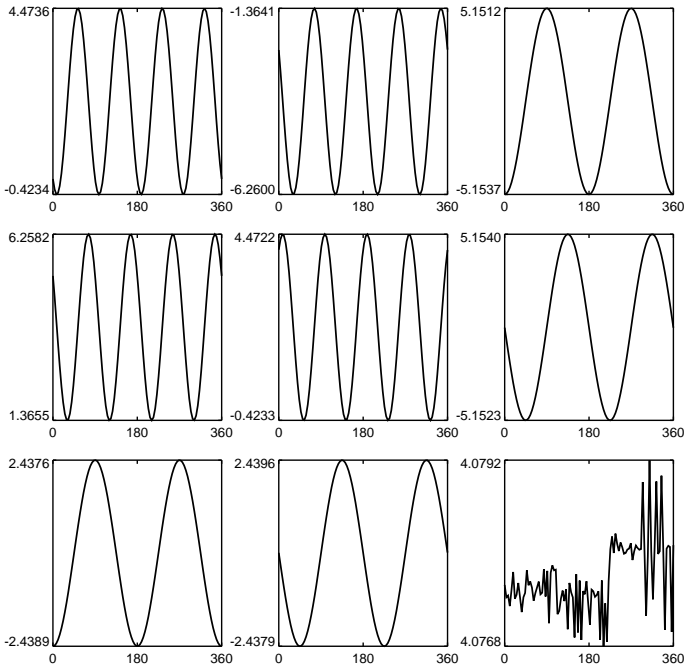


Fig. 8.— The difference of the Mueller matrices between a ten-layer achromatic $\lambda/4$ waveplate and an ideal $\lambda/4$ waveplate at a single wavelength $\lambda_0 = 1083.0$ nm. All the y-axis should be multiplied by 10^{-3} . And the x-axis is the azimuth of the waveplates. The first row and first column of the matrix are omitted, which are all zeros in the current numerical evaluation.

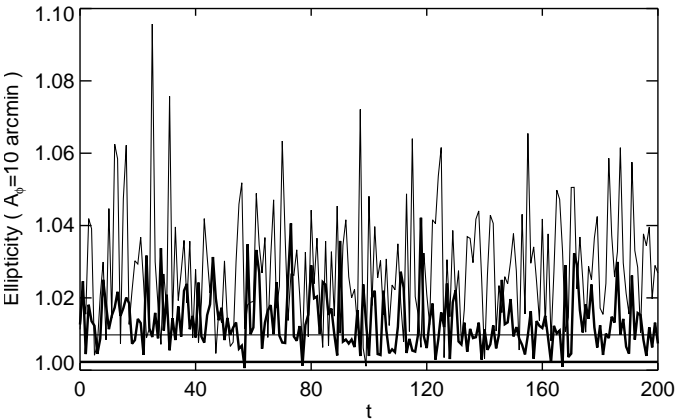


Fig. 9.— Performance stability of three- and ten-layer $\lambda/4$ waveplates at single wavelength $\lambda = 1523.1$ nm. Thick lines stand for three-layer waveplate; thin lines stand for ten-layer waveplate. Two straight lines stand for the situation when the random errors on each layer do not exist. The t-axis is an index of the random tests, not in real time units. The ellipticity of the testing light is obtained in the same way as in Figure 4 but for a single wavelength (see Section 4.2).

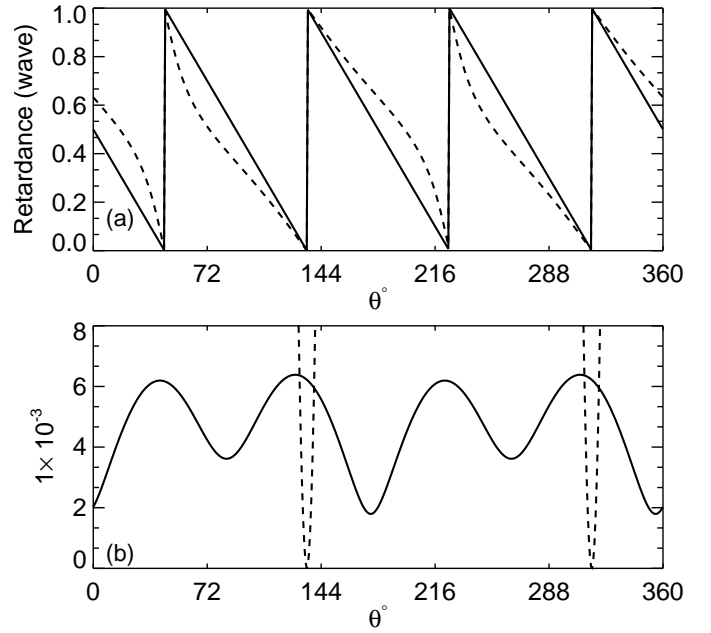


Fig. 10.— Phase shifter composed of three waveplates (see Section 5.1). The wavelength of the incident light is $\lambda = 1083.3$ nm. θ is the azimuth of the $\lambda/2$ waveplate in the center. The dashed curves refers to the corresponding properties of a phase shifter composed of a set of three waveplates made from single-layer uniaxial crystals, which are designed to be $\lambda/4$ or $\lambda/2$ waveplates at 1523.1 nm. The solid curves corresponds to a phase shifter made of three six-layer achromatic waveplates. (a) The phase change (retardance) vs. θ derived from the diagonal elements of Jones matrix and (b) the magnitudes of the anti-diagonal elements of the Jones matrices of the two phase shifters.

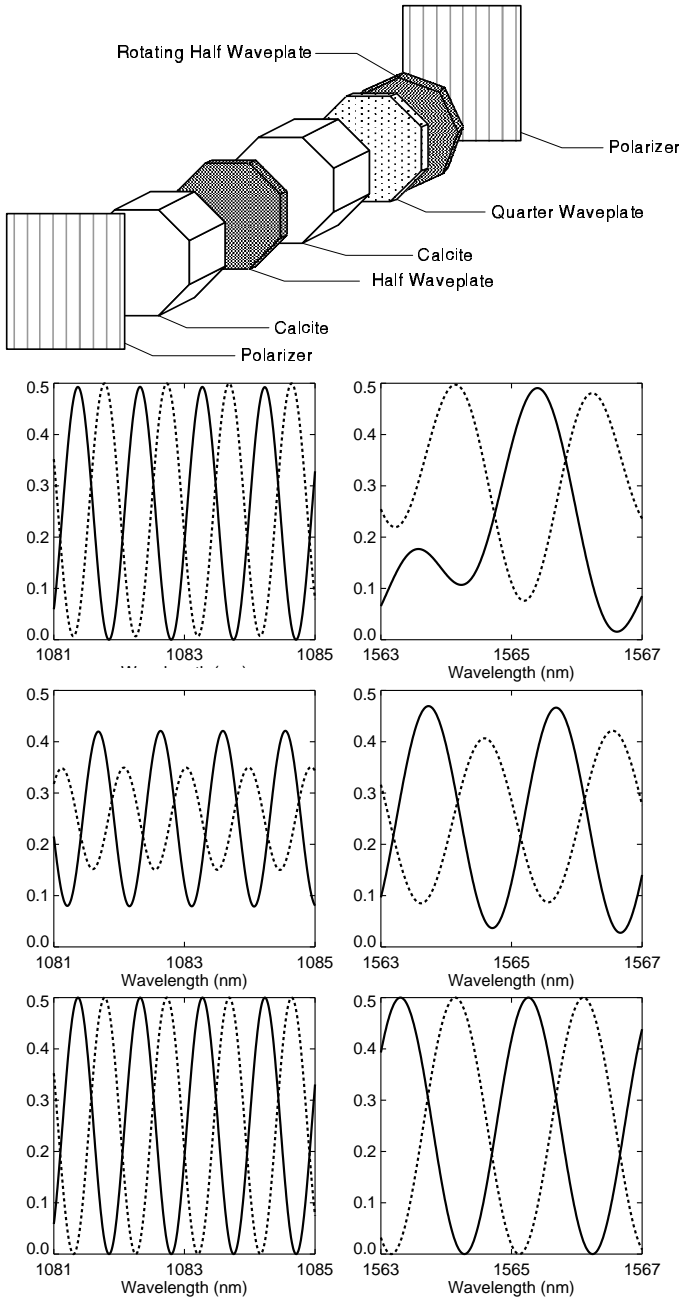


Fig. 11.— Tuning of a single stage of a Lyot filter - single stage. Solid curves: azimuth of the rotating $\lambda/2$ waveplate is set at 0; dotted curves: azimuth of the rotating $\lambda/2$ waveplate is set to $\frac{2}{7}\pi$. First row: a cartoon of a single stage; second row: transmission of the Lyot stage using single crystal waveplates; third row: using three-layer achromatic waveplates; fourth row: using six-layer achromatic waveplates.

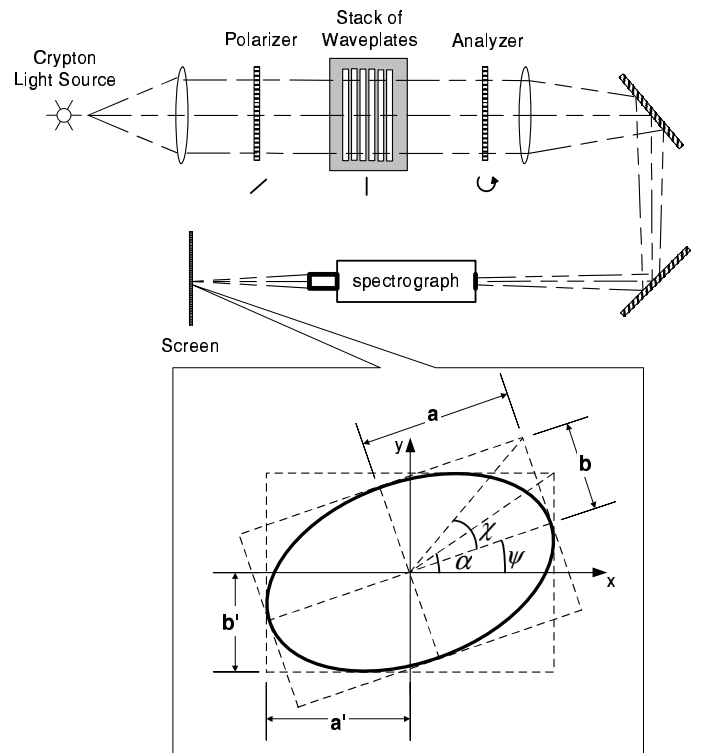


Fig. 12.— Measurement of the retardance. Shown in the waveplate-stack is the six-layer achromatic $\lambda/4$ -waveplate.

A. Jones Algebra and Mueller Algebra

In this study, we adopted the Jones matrix and Mueller matrix representations for all optical instruments and light sources. In this way, the description of the optical system can be put into a general framework and the designer is free of tedious vector combination and decomposition (cf., Evans 1949). The original formalization of these types of matrix algebra can be found in standard textbooks (for example, Born & Wolf 1999). For consistency, we list the matrix conventions used in the current paper.

The Jones matrix is a 2×2 matrix. A phase retarder can be written as the Jones matrix

$$\begin{bmatrix} e^{i\delta/2} & 0 \\ 0 & e^{-i\delta/2} \end{bmatrix}, \quad (\text{A1})$$

where δ is the retardance (phase difference) between the x - and y -components of the electric vector of a linearly polarized monochromatic wave after its passage through a retarder. For a linear polarizer (projection operator), the Jones matrix is

$$\begin{bmatrix} \cos^2 \theta & \sin \theta \cos \theta \\ \sin \theta \cos \theta & \sin^2 \theta \end{bmatrix}, \quad (\text{A2})$$

where θ is the active direction of the polarizer. These two optical instruments are frequently used in the current paper.

The Mueller matrix \mathbf{M} is a 4×4 matrix, which represents an optical instrument and can be derived from the corresponding Jones matrix \mathbf{J} according to

$$\mathbf{M} = [\mathbf{T} \cdot (\mathbf{J} \otimes \mathbf{J}^*) \cdot \mathbf{T}^{-1}]^T, \quad (\text{A3})$$

where

$$\mathbf{J} = \mathbf{R}(-\theta) \cdot \begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{-i\delta} \end{bmatrix} \cdot \mathbf{R}(\theta), \quad (\text{A4})$$

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix} \quad \text{and} \quad (\text{A5})$$

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{bmatrix}. \quad (\text{A6})$$

where “ \otimes ” denotes the Kronecker product, “ $*$ ” denotes the operation of complex conjugate, and “ T ” denotes transpose.

Furthermore, if we know that the vector representation of the light is the Jones vector \mathbf{L} (a complex column vector), then the Stokes vector \mathbf{S} is the vectorial form of the same light wave in Mueller algebra

$$\mathbf{S} = \mathbf{T}(\mathbf{L} \otimes \mathbf{L}^*). \quad (\text{A7})$$

The formalism of Jones algebra only describes the propagation of polarized light through an optical system. Typically, we encounter partial depolarization processes in an optical system. In these cases, Mueller algebra is more appropriate and reliable describing both intensity and polarization transformations. The proper choice entirely depends on the specific task in the optical design of the achromatic waveplate.

B. Statistical Simulation and Computing Methods

Simulated annealing algorithm is closely related to the Monte Carlo method used in statistical mechanics, specifically in condensed matter physics. The major task of using Monte Carlo method is to calculate the averaged quantities of a large size system, such as internal energy, specific heat, magnetization. We want to give a brief overview on the Monte Carlo method, then comment on the connection between the SA algorithm and Monte Carlo method. Eventually, we will have an idea about how the multiple-parameter optimization problem can be analogy to the physics problem in statistical mechanics in physics.

B.1. Monte Carlo Methods and Metropolis Algorithm

Monte Carlo method is used to calculate the expectation value of certain physical quantity, such as internal energy, specific heat, magnetization of the condensed matter. Let Q be such a quantity,

$$\langle Q \rangle = \frac{1}{Z} \sum_{\mu} Q \cdot e^{-\beta E_{\mu}} \quad (\text{B1})$$

where $\langle \cdot \rangle$ stands for the time average, Z is the partition function of the system defined as $Z = \sum_{\nu} e^{-\beta E_{\nu}}$; $\beta = \frac{1}{kT}$, k is the Boltzmann constant; μ or ν stands for the energy state corresponding to the system energy $E_{\mu, \nu}$. The difficulty of carrying out this summation is merely the large number of energy states of the system. What we can do is to somehow select a subset according to certain probability

p_μ of these energy states which contribute the most to the summation. This strategy is called *the importance sampling* and the result is an estimator of $\langle Q \rangle$:

$$Q_M = \frac{\sum_{i=1}^M Q_{\mu_i} \cdot p_{\mu_i}^{-1} \cdot e^{-\beta E_{\mu_i}}}{\sum_{i=1}^M Q_{\mu_i} \cdot p_{\mu_i}^{-1} \cdot e^{-\beta E_{\mu_i}}} \quad (\text{B2})$$

where a subset of the state space is selected as $\{\mu_1, \mu_2, \dots, \mu_M\}$. Now, if we choose $p_\mu = e^{-\beta E_\mu} / Z$, we will have

$$Q_M = \frac{\sum_{i=1}^M Q_{\mu_i}}{M} \quad (\text{B3})$$

It can be shown that this estimator is an excellent estimator if we are dealing the equilibrium system. The question is how we can find such a subspace of the state space?

Given a thermodynamic system, the energy distribution satisfies the following relation:

$$\frac{dw_\mu}{dt} = \sum_{\nu} [w_\nu \cdot P(\nu \rightarrow \mu) - w_\mu \cdot P(\mu \rightarrow \nu)] \quad (\text{B4})$$

where, w_μ is the occupation probability for the system staying at energy state E_μ , $P(\mu \rightarrow \nu)$ is the transition probability for the system moving from state E_μ to E_ν . This equation reveals the dynamics of the thermal system. As long as we know the dynamics of w_μ , we can derive basically everything of the system. In equilibrium, the left hand side of Equation B4 vanishes, i.e.,

$$w_\nu \cdot P(\nu \rightarrow \mu) = w_\mu \cdot P(\mu \rightarrow \nu) \quad (\text{B5})$$

By applying the summation over ν on both sides, and noticing that $\sum_{\nu} P(\mu \rightarrow \nu) = 1$, we have the following relation:

$$w_\mu = \sum_{\nu} w_\nu \cdot P(\nu \rightarrow \mu) \quad (\text{B6})$$

We will further write this equation into the matrix form as

$$\mathbf{w}(t+1) = \mathbf{P} \cdot \mathbf{w}(t) \quad (\text{B7})$$

where the matrix \mathbf{P} is called Markov matrix which describes the stochastic transition probability between any two energy states of the system. As $t \rightarrow \infty$, the system modelled with Equation B7 will eventually end up with two scenarios: (a) $\mathbf{w}(\infty) = \mathbf{P} \cdot \mathbf{w}(\infty)$, (b) $\mathbf{w}(\infty) = \mathbf{P}^n \cdot \mathbf{w}(\infty)$. The scenario (b) is the undesirable *dynamic equilibrium* (also called *limit circle*), in which the vector \mathbf{w} is changing length periodically. In order to eliminate this defect in our stochastic model, we put the following restriction

$$w_\mu \cdot P(\mu \rightarrow \nu) = w_\nu \cdot P(\nu \rightarrow \mu) \quad (\text{B8})$$

This relation is the most important relation in Monte Carlo method, called *detailed balance*. It can be shown the dynamic equilibrium and detailed balance are exclusive with each other. We also acknowledge that in scenario (a), \mathbf{w} will eventually overlap with the largest eigenvector of the Markov matrix \mathbf{P} .

According to Gibbs, in equilibrium, the energy distribution of a system associated with a thermal reservoir with temperature T will be the Boltzmann distribution,

$$w_\mu = \frac{1}{Z} e^{-\beta E_\mu} \quad (\text{B9})$$

Once we apply the Boltzmann distribution to the detailed balance in Equation B8, we have

$$\frac{P(\mu \rightarrow \nu)}{P(\nu \rightarrow \mu)} = \frac{w_\nu}{w_\mu} = e^{-\beta(E_\nu - E_\mu)} \quad (\text{B10})$$

In this equation, two things are included: (a) how to move to the next state ν from the current state μ ? (b) how to keep the ratio? It is not an easy task to satisfy both of these two requests with a simple choice of $P(\mu \rightarrow \nu)$. In practices of Monte Carlo methods, one efficient way to realize the process in Equation B10 is to split the $P(\mu \rightarrow \nu)$ into two parts,

$$P(\mu \rightarrow \nu) = g(\mu \rightarrow \nu) \cdot A(\mu \rightarrow \nu) \quad (\text{B11})$$

where g is the sampling probability, and the A is the acceptance probability (also called *acceptance ratio*). In so doing, we will have a greater flexibility of choosing g and A , only make sure that their values are between zero and one and adjust them to comply with the detailed balance relation. $g(\mu \rightarrow \nu)$ defines how to sample the state space given the current energy state μ . We can basically use any random process to do this, with one restriction — at least one path should exist between any two states in the state space. This is the *ergodicity condition* of the algorithm.

In practice, we are free to choose $g(\mu \rightarrow \nu) = g(\nu \rightarrow \mu)$. In so doing, we are left with only the acceptance ratio A to be decided:

$$\frac{A(\mu \rightarrow \nu)}{A(\nu \rightarrow \mu)} = e^{-\beta(E_\nu - E_\mu)} \quad (\text{B12})$$

One choice for $A(\mu \rightarrow \nu)$ is

$$A(\mu \rightarrow \nu) = A_0 e^{-\frac{1}{2}\beta(E_\nu - E_\mu)} \quad (\text{B13})$$

where A_0 is an adjustable factor to guarantee the acceptance ratio be in zero to one. This choice is a very small value for most transition. In order to use the CPU time efficiently, we want the acceptance ration to be as large as possible (the maximum is 1). An ideal scenario should be like this: the sampling process always generate new states exactly with the correct transition probability, and the acceptance ratio is always 1. In reality, the best we can do is to set the larger one between $A(\mu \rightarrow \nu)$ and $A(\nu \rightarrow \mu)$ to be 1, and adjust the other one according to the detailed balance.

$$A(\mu \rightarrow \nu) = e^{-\beta(E_\nu - E_\mu)}, \text{ if } E_\nu > E_\mu \quad (\text{B14})$$

$$A(\mu \rightarrow \nu) = 1, \text{ otherwise} \quad (\text{B15})$$

This is the Metropolis algorithm of handling the condition of the detailed balance. Monte Carlo methods based on Metropolis algorithm always accept the move towards lower energy states, and accept the move towards higher energy only according to the probability $e^{-\beta(E_\nu - E_\mu)}$.

The initial state of energy distribution of the system are always chosen to be either $T = 0$ or $T = \infty$. The reason is that the energy distribution of condensed matter under these two temperatures are always easy to calculate. In fact, in $T = 0$ case, the system is in the ground state, i.e., the Hamiltonian is in minimum. On the other hand, the energy states distribution of $T = \infty$ case is totally random. We are usually interested in the average of certain physical quantities at a given temperature different from 0 or ∞ .

Therefore, we can rephrase the Monte Carlo method for the condensed matter physics in the following way. We start the simulation with the energy states distribution corresponding to either $T = 0$ or $T = \infty$. Then the Markov process is programmed to generate a chain of states to simulate the transition process according to the restriction of the detailed balance, with which the resulting state distribution will be guaranteed to be the Boltzmann distribution because the detailed balance in Equation B10 is derived from the assumption that we will pick up the states according to the Boltzmann distribution. We then wait for the energy states distribution $w_\mu(t)$ to get close enough to the Boltzmann distribution, i.e., the equilibrium. The way to know that the system is at equilibrium is to monitor the value of the certain physical quantity at the current energy state accepted by the Monte Carlo program. In approaching the equilibrium, the value should turn more and more steady with only small amplitude fluctuation. Only after the system arrived in the equilibrium, can we start to calculate the physical quantities according to the Equation B3 while the Markov process is still running. The exact value of M is not important, increasing M will only improve the accuracy of the estimator.

B.2. Simulated Annealing

SA method is not necessarily a method dealing with any physics properties of a real physical system. SA method is only an algorithm used in solving the combinatorial optimization problem. We certainly can draw the analogy between general optimization problem and a condensed matter system in the perspective of mathematics, but we should not confuse the unique concepts in each of the two domains.

In SA method, the dynamics of the system, represented by a generalized energy function with free parameters to be decided, is put at a higher temperature (an effective temperature). Then the temperature is reduced according to an ‘‘annealing schedule’’. At each temperature the system experiences, a MC procedure is applied, i.e., the system is allowed to stay at the temperature long enough to go to an equilibrium, at the same time, the state space is sampled according to a sampling probability, and the final acceptance of the newly sampled state is decided from the acceptance probability. The decrease of the effective temperature continues until reaching zero (usually is a small number defined in computing program). This zero effective temperature corresponds to the ‘‘ground state’’ of the system, i.e., the minimum of the generalized energy function. Recall in Appendix B.1 that the ground state of a condensed matter system is almost always known and used as an initial point. The SA program used in the current paper is a most simplified routine in the sense of efficiency. And the routine will be updated in the future to handle more complicated computational problem in an efficient way.

In Table 3, we list several popular choices for sampling probability and the corresponding annealing schedule in simulated annealing.

C. Experiment On Waveplates

The scheme shown in Figure 12 is one of the method can be used in measuring the retardance of waveplate, assuming the optical axis of the waveplate is known before hand. The output intensity spectrum is recorded at the screen, which is usually a CCD sensor array. Since the light after passing through the polarizer and a $\lambda/4$ is elliptically polarizer due to the imperfectness of the waveplate, the output intensity varies with the azimuth of the analyzer periodically. The maximum and minimum of the intensity can therefore be selected, which are proportional to a and b respectively in Figure 12. The ratio of a/b is the ellipticity shown in Figure 4. Notice that another commonly used measure of the elliptical shape, eccentricity ε , is related to the ellipticity by $\varepsilon = \sqrt{1 - b^2/a^2}$. Now we want to show how the retardance δ of the waveplate can be derived from the ellipticity and the incline angle ψ .

The following relations between the angles in Figure 12 can be easily obtained (see, for example, Born & Wolf 1999):

$$\sin 2\chi = \sin 2\alpha \sin \delta \quad (\text{C1})$$

$$\tan 2\psi = \tan 2\alpha \cos \delta \quad (\text{C2})$$

from which, we can derive

$$\sin 2\alpha = \sqrt{(\sin^2 2\chi + \tan^2 2\psi) / (1 + \tan^2 2\psi)} \quad (\text{C3})$$

where, $0 \leq \alpha \leq \pi/2$ by definition. In this relation, ψ can be measured in the experiment, which is the azimuth of the analyzer while the output intensity is maximum. Also,

$$\tan \chi = \mp b/a \quad (\text{C4})$$

which is related directly to the measurable ellipticity. Therefore, α angle can be found, and so can the retardance δ . Moreover, a nontrivial caution should be taken due to the following complications:

$$0 < \psi < \pi/2 \quad \text{if} \quad 0 < \delta < \pi/2 \quad (\text{C5})$$

$$\pi/2 < \psi < \pi \quad \text{if} \quad \pi/2 < \delta < \pi \quad (\text{C6})$$

In case of using Equation C1 to derive δ , we take

$$\delta = \sin^{-1} \left(\frac{\sin 2\chi}{\sin 2\alpha} \right), \text{if } 0 < \psi < \pi/2 \quad (\text{C7})$$

$$\delta = \pi - \sin^{-1} \left(\frac{\sin 2\chi}{\sin 2\alpha} \right), \text{if } \pi/2 < \psi < \pi \quad (\text{C8})$$

These two relations were implemented in deriving the plots in Figure 5.

TABLE 3
EXAMPLES OF SIMULATED ANNEALING ALGORITHM

	Sampling Probability	Annealing Schedule	Category	Acceptance Probability (Metropolis)
Gaussian Distribution	$(\frac{1}{2\pi T})^{D/2} \exp(-\frac{\Delta x^2}{2T})$	$T(t) = T_0 / \ln t$	Annealing	
		$T_{t+1} = T_t - T_0 \ln t_0 / [t(\ln t)^2], t \gg 1$	Annealing	
		$T_{t+1} = cT_t, 0 < c < 1$	Quenching	Same for all:
		$T_{t+1} = T_t [1 + (c+1)\Delta t], t \gg 1$	Quenching	$\min[\exp(-\Delta E/T), 1]$
		$T_t = T_0 \exp[(c-1)t]$	Quenching	
Cauchy Distribution	$\frac{T}{(\Delta x^2 + T^2)^{(D+1)/2}}$	$T(t) = T_0/t$	Fast Annealing	

NOTE.— Δx stands for the path difference between two adjacent states. T is the effective temperature in simulated annealing process. t is the measure of time in Monte Carlo steps. D is the dimension of the state space. ΔE is the energy difference corresponding to the states in calculating Δx as long as the newly sampled state is accepted according to sampling probability, (See Ingber 1996, for more details).