ABSTRACT

DESIGN AND APPLICATION OF NEAR INFRARED TUNABLE FILTER FOR ATST AND NST

by

Jun Ma

This thesis includes the following sections: a general design of the Near Infrared Tunable Filter (NIRTF) of Advanced Technology Solar Telescope (ATST) project and the Near Infrared Imaging Magnetograph (IRIM) of the New Solar Telescope (NST), the design of achromatic waveplates, the narrowband continuum observation of a Sunspot at 1.56 μm using IRIM instruments, and the coronal hole observation using the Digital Vector Magnetograph (VMG) at Big Bear Solar Observatory (BBSO).

A near infrared tunable filter system is designed with detailed optical parameters for each component in the system. Among these components, the achromatic waveplates, which used to be an obstacle to the design of tunable filter systems, are redesigned using a Monte Carlo global optimization algorithm. The designs for three sets of such multi-layer achromatic waveplates are presented in the thesis. Experimental results are also obtained for three-layer quarter and half waveplates. These results show that the chromatism of waveplates is reduced to a satisfying level for the filter system in the wavelength range 1000$\sim$1800 nm.

A set of narrow band continuum images at 1.56 μm are analyzed for the active region NOAA 10707 using the Lyot filter in the Near Infrared Imaging Magnetograph (IRIM) system and high order adaptive optics at National Solar Observatory (NSO), New Mexico. The analysis shows that the horizontal flow field within umbral region is small, in the order of 0.3 km$\cdot$s^{-1} on average. Also, a narrow ring (370 km) within which the proper motion of the flow in the sunspot reverses its direction from inward to outward is identified. For the matured sunspot like the one in NOAA 10707, a gradual shrinking of spot is found, which suggests that at least in the early stage of the decay of matured active regions, a reduction of area takes place, instead of decomposing immediately.

A preliminary analysis of the magnetic flux data using the DVMG at BBSO reveals that, on the photospherical level, the magnetic flux evolves faster in quiet sun regions than in the coronal hole. Also fewer bipolar features are connected by arch filament (shown in Hα images) in the coronal hole.In the coronal hole, most of the closed magnetic loops appear at level lower than 5000 km, while open magnetic loops fill the space above 5000 km level. This observation suggests that flux cancellation or magnetic reconnection in photosphere and lower chromosphere may only take place at an atmospheric level lower than 5000 km, instead of the height where fast solar wind is believed to generate (between 5000$\sim$20000 km).

The science deduced based on these observations could be greatly improved by obtaining accurate measurements of the magnetic field strength (instead of flux density) and filling factor, which are the primary goals of the newly designed Near Infrared Tunable Filter system (NIRTF) in this thesis for the ATST and the New Solar Telescope (NST).

DESIGN AND APPLICATION OF NEAR INFRARED TUNABLE FILTER FOR ATST AND NST

by
Jun Ma

A Dissertation
Submitted to the Faculty of
New Jersey Institute of Technology and
Rutgers, the State University of New Jersey - Newark
in Partial Fulfillment of the Requirements for the Degree of
Doctor of Philosophy in Applied Physics

Federated Physics Department

August 2007

APPROVAL PAGE

DESIGN AND APPLICATION OF NEAR INFRARED TUNABLE FILTER FOR ATST AND NST

Jun Ma

Dr. Haimin Wang, Dissertation Co-Advisor — Date
Distinguished Professor of Physics, Associate Director of the Center for
Solar-Terrestrial Research and Big Bear Solar Observatory, NJIT

Dr. Carsten Denker, Dissertation Co-Advisor — Date
Assistant Professor of Physics, NJIT

Dr. Philip R. Goode, Committee Member — Date
Distinguished Professor of Physics, Director of the Center for
Solar-Terrestrial Research and Big Bear Solar Observatory, NJIT

Dr. Dale E. Gary, Committee Member — Date
Professor of Physics, Director of the Solar Array
in Owens Valley Radio Observatory, NJIT

Dr. Zhen Wu, Committee Member — Date
Professor of Physics, Rutgers University, Newark

BIOGRAPHICAL SKETCH

Author: Jun Ma

Degree: Doctor of Philosophy

Date: August 2007

Undergraduate and Graduate Education:

- Doctor of Philosophy in Applied Physics,
  New Jersey Institute of Technology, Newark, New Jersey, 2007
- Master of Science in Mechanical Engineering,
  New Jersey Institute of Technology, Newark, New Jersey, 2002
- Bachelor of Science in Aerodynamics & Flight Mechanics,
  Nanjing University of Aeronautics & Astronautics, China, 1998

Major: Applied Physics

Publications:

Ma, J., Denker, C., & Wang, H., *Optical Design of Multilayer Achromatic Waveplates by Simulated Annealing and Their Applications*, 2007, submitted to PASP.

Zhang, J., **Ma, J.**, & Wang, H., 2006, Astrophys. J., 649, 464.

Cao, W., Jing, J., **Ma, J.**, Xu, Y., Wang, H., & Goode, P. R. 2006, PASP, 118, 838.

Cao, W., Hartkorn, K., **Ma, J.**, Xu, Y., Spirock, T., Wang, H., & Goode, P. R. 2006, Sol. Phys., 238, 207.

Xu, Y., Cao, W., **Ma, J.**, Hartkorn, K., Jing, J., Denker, C., & Wang, H. 2005, Astrophys. J. Lett., 628, L167.

Ma, J., Wang, J., Cao, W., Denker, C., & Wang, H. 2004, Proc. SPIE, 5523, 139.

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., Didkovsky, L., **Ma, J.**, Shumko, S., Varsik, J., Wang, J., Wang, H., & Goode, P. R. 2003, Astronomische Nachrichten, 324, 332.

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.

To My Beloved Parents

ACKNOWLEDGMENT

I am very grateful to my advisor, Dr. Haimin Wang, for his mentorship, without which it would have been impossible for me to complete my thesis. I also want to thank my co-advisor, Dr. Carsten Denker, for his guidance throughout my research and for showing me how to explore the natural areas around Big Bear Lake. I extend my gratitude to the other committee members for their tremendous support on my dissertation: Dr. Philip Goode, Dr. Dale Gary, and Dr. Zhen Wu.

Special thanks to the scientific staff and operational staff at Big Bear Solar Observatory (BBSO) in California, for their kindness and assistance during my stay. I want to thank Dr. John Varsik, Sergiy Shumko, Dr. Vasyl Yurchyshyn, Dr. Valentyna Abramenko, Dr. Wenda Cao, Randy Fear, and Jeff Nenow for their encouragement and advice. In particular, Dr. Thomas Spirock for his help with the magnetograph and many other instruments at the BBSO facilities. I also enjoyed the time I rode in his Rubicon wondering around even though we never had a chance to conquer the John Bull Trail.

I wish to express appreciation to Dr. Jingshan Wang for sharing his expertise on filter designs and providing valuable suggestions on my dissertation.

My sincere gratitude goes out to Dr. Jun Zhang of Beijing Astronomical Observatory in China for being a source of inspiration. I was intrigued by our discussions on magnetic fields distribution, which eventually led to one part of my dissertation. His extensive scientific knowledge and techniques broadened my horizons.

Finally, I want to thank Dr. Weijun Mao of Nanjing Institute of Astronomical Instrumentation (China) and his staff for manufacturing and testing the waveplates discussed in my thesis.

This thesis work was supported by NSF under grants ATM 03-42560, ATM 02-36945, IIS ITR 03-24816, AST MRI 00-79482, and by NASA under grant NNG0-6GC81G.

TABLE OF CONTENTS

Chapter **Page**

TABLE OF CONTENTS

(Continued)

Chapter **Page**

TABLE OF CONTENTS

(Continued)

Chapter **Page**

TABLE OF CONTENTS

(Continued)

Chapter **Page**

LIST OF TABLES

Table **Page**

LIST OF FIGURES

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

LIST OF FIGURES
(Continued)

Figure **Page**

CHAPTER 1

INTRODUCTION

The Sun is the only star that is close enough for scientists to observe in detail to discover the fine structures. The purpose of these scientific observations are beyond the matter of curiosity, since the radiation from the Sun not only provides the only energy source for all lives on the Earth, also it can produce hazardous side effects, for example, space storms induced by solar eruptions. By recording various solar activities and finding proper explanations of the driving mechanisms for these activities, it will be possible to predict future eruption events on the Sun. So that preparations can be made for those events on the Sun that are hazardous to the human activities on the Earth and in space.

The activities on the Sun are all related to the magnetic fields, which are believed to be generated by a solar magnetic dynamo (Cowling 1981): electric currents tend to form while the conducting fluid is moving through an external magnetic field; the generated currents will produce another magnetic field superimposed on the external magnetic field and redefine the total magnetic field in space. The dynamics of solar magnetic field are described by the induction equation (Stix 2002):

$$\dot{\mathbf{B}} = \nabla \times (\mathbf{v} \times \mathbf{B}) - \nabla \times (\eta \nabla \times \mathbf{B}) \tag{1.1}$$

where, $\mathbf{v}$ is the velocity field of plasma; η is the magnetic diffusivity. In solar plasma, the induction term (the first term on the right side of equation) outweighs the second term which represents the Ohmic dissipation. The velocity fields needed to maintain the dynamo process are provided by convection and differential rotation (Solanki 2003). The flux ropes generated by the solar dynamo are brought up to the surface of the Sun by convective motions, and manifest themselves as various visible features (Weiss et al. 1996; Schüssler & Vögler 2006; Meyer et al. 1974), for example, sunspots, pores and networks.

Solar observations focus on dthe measurement of the properties of these features and their evolutions, based upon which more precise physical models can be constructed. Large magnetic features, i.e., sunspots, have been under extensive research for decades (Parker 1974, 1979; Rimmele & Marino 2006; Thomas et al. 2002). However, the fine structure of sunspots is always under debate. Among these unsettled arguments are: what are umbral dots? And what are the physical properties of plasma occupying the darkest regions in the sunspot? The answers to these questions are the keys towards understanding the nature of sunspots, and subjects related to global magnetic field evolution on the Sun (Howard & Labonte 1981; Solanki et al. 2000, 2006).

On the other hand, weak magnetic features, such as those appearing in quiet sun regions and coronal hole regions, reveal a different view of the solar magnetic activity. It was recognized that the magnetic flux embedded in these regions contributes the principle fraction of the total solar flux (Domínguez Cerdeña et al. 2006), and they are magnetic phenomena that are independent of the global solar dynamo (Petrovay & Szakaly 1993; Okunev et al. 2005). Furthermore, the photospheric distribution of the flux and field within coronal holes is correlated with many important questions, such as the origins of solar wind (Tu et al. 2005).

Better understanding of these subjects relies primarily on the improvements of instrumentations used in observation. The greatest improvements in techniques in recent years are the development of near infrared detectors (Ives & Bezawada 2007), narrow band tunable filters (Wang et al. 2001; Tritschler et al. 2002; Denker et al. 2003b,a), adaptive optics (Rimmele 2000; Rimmele et al. 2003) and new, large aperture ground-based telescopes, such as 1.5 m GREGOR (Volkmer et al. 2006), 1.6 m New Solar Telescope (NST) (Didkovsky et al. 2003, 2004), 4 m Advanced Technology Solar Telescope (ATST) (Rimmele et al. 2004, 2005).

ATST 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 0.3 to 28 μm (see on-line ATST project documents at: http://atst.nso.edu/library/specs.shtml) with minimal scattered light and high order adaptive optics. Several major focal plane instruments are currently being designed to exploit the high resolution capabilities of ATST. The Near Infrared Tunable Filter (NIRTF) system is one of these proposed focal plane instruments, which is designed at Center for Solar-Terrestrial Research at New Jersey Institute of Technology (NJIT).

Big Bear Solar Observatory (BBSO) is building an 1.6-meter off-axis telescope as an upgrading of its current 65 cm telescope. A preliminary tunable filter system which has narrower infrared spectrum coverage, will be integrated into the focal plane instrumentation of NST as well.

In the following a few subsections, more detailed background research are presented regarding the subjects mentioned above.

1.1 Sunspot Structure

Several photometric properties of sunspots are of interests: sunspot size variation, penumbral flows, umbral topology, penumbra-umbra interaction. These properties directly relate to the energy transport within active region. Sunspot size variation is related to the decaying process of sunspots, i.e., the loss of magnetic flux due to the interaction with the turbulent plasma motions around the sunspot. The heat flux loss is vividly represented by the penumbral grains in the model of Evershed flow, in which the penumbral filaments are bent parallel to the solar surface becoming diffuse at one end (Rimmele & Marino 2006). One explanation for the cooling of the sunspot is that the kinetic energy carried by the out-going penumbral filaments balance or partially balance the energy loss within sunspot. The most detailed structures in umbra are the umbral dots (UD). These dots have excessive brightness with respect to the surrounding umbral region, and possess motions similar to granulation cells. However, in this case the strong magnetic fields (1500$\sim$3000 G) surrounding the dots suppress the convection motions.

Two sunspot models have been proposed in the literature decades ago: the cluster model (Parker 1979), and the monolithic flux-tube model (Cowling 1975; Meyer et al. 1977). Despite years of extensive observations and theoretical calculations, neither of them can be dismissed confidently. As pointed out by Parker (1979), the explanation of umbral dots should be accounted for to validate the models. The cluster model interprets UD as convective overstability of field-free gas columns. However, this model could not explain why these penetrating gas columns appear as isolated dots instead of forming bright network along the boundaries of bundles of flux tubes. The monolithic flux-tube model describes the UDs as the overshooting of coherent cellular motions of a length scale of 250$\sim$300 km extending over a depth of 1500 km below the visible photospheric layer (Knobloch & Weiss 1983). The same argument also accounted for the observed umbral oscillation and UD oscillation. In addition, the correlation between chromospheric UDs and photospheric UDs was predicted in Zirin (1974); Loughhead (1974). By comparing the observations at Hα (6563 Å) and G-band (4308 Å), Kitai (1986) proposed that UDs are not of the same convection origin as granules (Bumba & Suda 1980). The correlation between chromospheric UDs and photospheric UDs suggested that UDs are of some different and unique formation mechanism, since it is difficult to accept that the overturning motions are able to shoot up to $\tau_{H\alpha} = 1$ level. Since the apparent difference between the two models lies in the topological structures of active regions, high resolution observations at different wavelengthes are critical to discriminate the two models.

The umbra of a sunspot is a highly inhomogeneous region considering the photometric intensity and magnetic field distribution (Severnyi 1965; Abdusamatov & Krat 1969) with-in the region. It is now widely accepted that an umbral region is composed of several subareas according to their brightness and morphology, namely, umbral cores (UCs), light bridges (LBs), and umbral dots (UDs). UCs are the darker regions in a sunspot, occupied by very strong magnetic fields. Convective motions in UCs are believed to be suppressed by such strong magnetic field. Hence the energy emission in UCs is highly reduced and the

regions appear to be darker. UDs are isolated bright dots embedded in UCs background and their size (equivalent diameter) ranges from 0".14 up to almost 1", with life time spanning from several minutes to hours (for example, see the overview by Sobotka et al. (1997a) and Sobotka et al. (1997b)). The discussion of the convective nature of UDs is also found in Schüssler & Baumann (2006). LBs are stream-like bright features penetrating the umbra and are highly stable. High spatial resolution observations reveal that a LB is usually composed of a sequence of well aligned bright dots which are similar to UDs (Sobotka et al. 1994).

Sobotka et al. (1992) found that the apparent intensities of a single UD and its corresponding umbral background appear to be highly correlated based on an analysis of 29 well-resolved bright features. UDs have not been reported within darkest UC regions. Accordingly, they concluded that the stronger the magnetic field is, the deeper the convection will take place. In one of their later papers (Sobotka et al. 1992), this correlation was confirmed again based on the analysis of three bright features, including clusters of UDs. Wiehr (1994) found that the contrast of UDs with respect to the local diffuse background decreases with increasing geometrical height. This discovery leads to a physical picture in which the convection flux diminishes with the increase of geometrical height.

Through high resolution spectra analysis, Lites et al. (1991) found that the background magnetic field strength in an umbra varies between 1400~2400 Gauss, without significant reduction by the appearance of central UDs, which is the opposite to the observed decreasing flux in LBs. Tritschler & Schmidt (1997) also observed that both central UDs and peripheral UDs (dots near the boundary of umbra) are embedded on the background magnetic field without significantly reduce the field locally.

One way to validate the convective nature of UDs is the vector velocity field. The diffraction-limited Doppler filtergram obtained by Rimmele (2004) shows that the velocity of vertical up flows is about 1 $km \cdot s^{-1}$ in UDs, which is on the same order as the theoretical value predicated by the cluster model of sunspot Parker (1979). The central UDs are almost

always static. The peripheral UDs are hardly discriminated from bright spots broken down from the penumbral grains, which drift inward. Rimmele (2004) suggested that central UDs and peripheral UDs can be distinguished from each other by observing their proper motion.

1.2 Photospheric Observation of Coronal Hole Magnetic Field

Solar magnetic field flux density varies largely from 10 Gauss in intra-network up to $\sim$3000 Gauss in active regions. One solar magnetic field model shows that the magnetic flux are buried under the convection zone in quiet regions. The disturbance by supergranular convection brings the magnetic flux up to the visible surface and forms a great variety of magnetic features: sunspots, pores, faculae, networks and intranetworks (IN). Due to the solenoidally of the magnetic field, the field lines often form closed loops. When both polarities of a loop locate on the solar surface, the loop structure can be clearly seen from, for example, H-α 656.3 nm chromosphere images or soft X-ray images from even higher up atmosphere. The field lines could also stretch into the interplanetary space, and form open field line regions leaving only one polarity on the solar surface.

Coronal holes are low density regions, in which field lines are open to interplanetary space. High energy particles escape from coronal holes easily and turn out to be the major contributor of the solar wind. Polar coronal holes usually have long lifetimes, about 8 years and are correlated to the 11-year periodic global magnetic field reversal on the Sun. The lifetime of nonpolar coronal holes are measured in several solar rotations and are related to the relics of diffused active regions. Short-lived coronal holes are related to coronal mass ejections (CMEs) and last for only a few days (Harvey & Recely 2002). The magnetic field characteristics of coronal holes on the photosphere provide the first step to the understanding of the mechanism of coronal holes.

1.3 Spectropolarimetry and Instrumentation

Solar observations are usually carried out by utilizing polarimetry instruments, such as spectrographs, Lyot filters, variable retarders and polarizers (Stix 2002). Filter-based spectropolarimetry instruments use narrow band filters to obtain 2D images, for example, the Universal Birefringent Filter System (UBF) of NSO (Beckers et al. 1975), Visible Imaging Magnetograph (VIM) (Denker et al. 2003a) and Infrared Imaging Magnetograph (IRIM) (Denker et al. 2003b) of BBSO. Spectrograph-based instruments use a spectrograph instead of filters to obtain 1D images at multiple spectral lines simultaneously, as in the Zurich Imaging Polarimeter (ZIMPOL) (Gandorfer 1999) for example. Since the narrow band tunable filters, such as Fabry-Perot etalon and Lyot filter, become more commonly used, the tunable-filter-based instruments have become more favorable for solar observation. The obvious advantages of the tunable-filter-based instruments are: larger field of view (FOV), direct 2D observation, and near-simultaneous multiple-line observation (due to rapid wavelength tuning and reduction of other instrumental delays).

The basic principle of these instruments is spectropolarimetry, which is an extremely powerful tool in the measurement of the magnetohydrodynamic motion and features on the Sun. Most of the important physical parameters of the solar atmosphere, such as chemical composition, density, and magnetic field, can be inferred from the spectropolarimetry observation. Not only is this information of vast interest to the empirical research of the solar activity, but also it provides the theoretical astrophysics guidelines and constraints to the magnetohydrodynamic models.

Primarily, spectropolarimetry information concerns the measurement of the deformation of spectral lines in solar spectrum. Magnetic-sensitive spectral lines are split in the presence of magnetic field permeating certain regions in the solar atmosphere. Due to the limited magnetic field strength in most area of the Sun, usually the splitting is small compared to the intrinsic line-broadenings due to temperature (Gaussian broadening) and pressure (Doppler broadening). And instead of a clear splitting, they just show additional

line broadening. A simplified model of such splitting is the Zeeman triplet model: two σ-components being circularly polarized, and one π-component being linearly polarized. In this scheme, the shifting between the two σ components is proportional to the magnetic field strength that induces the splitting. Only within regions supporting strong magnetic field (for examples, sunspots and pores), can the split profiles of triplet components be distinguished. When saturation happens, it usually implies that the σ components are shifted so far apart that one of them or both of them actually move out of the bandpass of the filters used in observation. Under the weak field assumption, Zeeman splitting can be written as (Strous 1994):

$$\Delta\lambda_z = \frac{e}{4\pi m_e c^2}\lambda^2 gB = 18.52\left(\frac{\lambda}{6302}\right)^2 gB \ \text{m/kG} \tag{1.2}$$

In order to detect weaker fields, longer wavelengths with greater Landé-g factors are more desirable.

The polarimetric properties of the Zeeman triplet model make it possible to distinguish each component using optics such as phase shifters and polarizers. Such an approach has been a major focus of solar spectropolarimetry for decades. Two different tracks were followed: spectrograph based and imaging based. The basic idea of the latter is to select the spectral line of interest using a prefilter, which has a pass band broad enough to include the overall split profile of the line. Then the filtered light passes through a polarization analyzer, and each component of the Zeeman triplet can be selected sequentially. The final product of this procedure is an array of images of every component of the triplet for a region of interest (ROI). The spectrograph-based approach utilizes a high spectral resolution spectrograph to scan the ROI, which can cover many spectral lines at a time. However, the instantaneous size (field of view) of the ROI is severely limited by the slit of spectrograph, which is then scanned spatial across the Sun to enlarge the ROI. This compromise comes from the basic principle behind any spectrograph. The spectrograph-based approach can also (usually does) work with a polarization analyzer in order to measure the detailed polarization signals.

There are several types of high spectral resolution filters available: Fabry-Perót interferometer (FPI), and Lyot filter. Compared to the high precision, high resolution, but mechanically fragile FPI, the Lyot filter has a more robust mechanical structure, and fairly high resolution, yet also has a relatively low transmittance that will reduce the contrast of features. Instruments based on the Lyot filter are the primary focus of the current thesis. Other types of high resolution filter, such as Michelson interferometer, also have applications in filter based solar observation tasks.

The Lyot filter, sometimes called Lyot-Ohme filter, was named after the French scientists Lyot and Ohme. In principle, it is a birefringent filter, i.e., utilizing the birefringence of a uniaxial crystal to achieve a narrow spectral bandpass. The term uniaxial crystal refers to a crystal which has a preferred direction, i.e. the extraordinary axis. An electromagnetic wave whose amplitude vector is parallel to this direction travels faster (or slower for negative crystal) than any other directions within the crystal. The difference between the two travelling speeds introduces a phase delay δ:

$$\delta = \frac{\mu d}{\lambda} 2\pi \tag{1.3}$$

where, μ is the dimensionless birefringence, d is the travelling distance in the crystal, and λ is the wavelength of wave. In later sections, the birefringence will be discussed in details. In short, on one hand, the wavelength-dependent birefringence is exactly the property needed for the crystal in a Lyot filter to disperse the spectrum in wavelength; on the other hand, the phase delay δ of some optical components in the filter are preferred to be independent of wavelength, i.e., achromatic.

1.4 Achromatic Waveplate Design and Algorithm

The retardance produced by an achromatic waveplate must be independent of the wavelength of the light passing through. For example, an achromatic quarter waveplate in-

troduces quarter-wave ($\pi/2$) phase difference between the two components of light ray regardless of the wavelength of the light.

An achromatic waveplate can be obtained by placing thin layers of birefringent crystals with properly designed thicknesses and azimuthal angles (Pancharatnam 1947). The design problem is therefore an optimization problem to minimize the chromatism by manipulating these free parameters, i.e., thicknesses and azimuths. The analysis method applied in the current thesis is Simulated Annealing (SA), which belongs to the category of stochastic optimization algorithms.

The underlying mathematical principle of the SA is the Markov chain. This specific stochastic process has many unique properties. One of the most important properties is that an irreducible Markov chain will converge to an equilibrium distribution (see later sections for details) (Otten & van Ginneken 1989). In the physical world, this property corresponds to the annealing of melted material (Kirkpatrick et al. 1983) — no matter how high the initial temperature was, as long as the annealing is conducted with certain constraints, the final inner structure will be arranged in a regular fashion, i.e., crystalline structure.

CHAPTER 2

FILTER SYSTEMS AND INFRARED IMAGING MAGNETOGRAPH

2.1 Principles of Solar Spectrography

The principles of solar spectral-polarimetry lie in radiative transfer (Unno 1956; Stix 2002). The polarization of sunlight carries information of the solar atmosphere through which the radiation passed. The polarization of the radiation is usually described by a Stokes vector: (I,Q,U,V), a vector independent of any specific coordinate system. The I component carries the information of the total energy flux of the radiation; Q,U components carry linear polarization energy flux (along two directions of a relative angle 45°); and V component bears circular polarization energy flux.

The Stokes V component is of particular importance, not only because it is related to the line-of-sight magnetic field $B_{\parallel}$, but also because it is the component that can be measured with higher accuracy than Q,U components. Stokes V can be written as (Stix 2002):

$$\begin{aligned} V(\lambda) &= \frac{1}{2}[I(\lambda+\Delta\lambda_z)-I(\lambda-\Delta\lambda_z)] \\ &= \Delta\lambda_z\left[\frac{\partial I}{\partial\lambda}+\frac{1}{6}(\Delta\lambda)^2\frac{\partial^3 I}{\partial\lambda^3}+\cdots\right] \end{aligned} \tag{2.1}$$

where, $\Delta\lambda_z$ is the Zeeman splitting. Given the assumption of weak field, omit higher order terms and consider Zeeman splitting relation $\Delta\lambda_z \propto B\cos\gamma$, so that

$$B\cos\gamma = \frac{c_1}{g}\cdot\frac{V}{\partial I/\partial\lambda} \tag{2.2}$$

where, γ is the inclination angle of the field vector, c_1 is a calibration factor and it is a function of wavelength λ, g is effective Landé factor. Although usually, more sophisticated

inversion routines are used to find magnetic field strength from the observed Stokes signals along with other parameters such as filling factor and the flux of the field (see for examples, Socas-Navarro et al. (2000); Socas-Navarro (2001)). A similar relation can be derived for Q, U as well for parallel field (Stix 2002):

$$B \sin\gamma = \frac{c_2}{g} \cdot \left[\frac{\Delta\lambda \sqrt{Q^2+U^2}}{\partial I/\partial\lambda} \right]^{\frac{1}{2}} \tag{2.3}$$

where, c_2 is a calibration factor; $\Delta\lambda$ is the offset from line center in Å(Mickey et al. 1996); g is the effective Landé factor.

The Stokes components are usually measured in combinations. Given two variable retarders, for example liquid crystal variable retarder, VR_1 and VR_2, let their retardance be δ_1 and δ_2. Usually the retardance is a function of the applied voltage on the variable retarder. Further, assume that the relative angle between the optical axes of the two variable retarders is 45°. Then the measurable of the incoming light is

$$I = (I_0 + Q_0 \cos\delta_2 + U_0 \sin\delta_1 \sin\delta_2 + V_0 \cos\delta_1 \sin\delta_2)/2 \tag{2.4}$$

where, I is the intensity measurable, I_0, Q_0, U_0, V_0 are Stokes components of the sunlight. Therefore, pairs of δ_1 and δ_2 can give different combinations of the components of Stokes vectors. For example, $\delta_1 = \lambda/4, \delta_2 = -\lambda/4$, the measurable is $\frac{1}{2}(I_0 - U_0)$. Equation 2.4 also can be used to measure the retardance of a retarder, given that one of the two retarder is an accurately calibrated retarder — for example, a waveplate — and the incoming light (I_0, Q_0, U_0, V_0) is also known.

2.2 NIRTF

The NIRTF is designed for the wavelength range from 1.0~1.7μm according to the scientific requirement of ATST (Rimmele et al. 2004, 2005; Keil et al. 2004). The main chal-

lenges of designing this filter are: wide wavelength coverage, large field of view (FOV), narrow bandpass and tunability. A summary of these scientific requirement of NIRTF is listed in Table 2.2.

The NIRTF consists four optical components: prefilter, polarization analyzer, narrow band filter, and NIR imaging array. Prefilter can be interference filters, which utilize coating techniques to bring several thin films together to form stratified medium and therefore interference. The interference pattern is a function of wavelength and usually is broad. The limitation of the interference filter is that the coating technique can only serve for a certain wavelength range only. If the two wavelength ranges are far apart, then another interference filter has to be designed. The prefilters needed for NIRTF are: for 15648Åand 15652Å lines, the bandpass of the filter is 15650Å±25Å; for 10830Å line, 10830±15Å.

Polarization analyzer can be made of liquid crystal. The functionality of polarization analyzer is described in Equation 2.4 — to select polarization components of the light. LOS field measurements require a modulator composed of a near infrared $\frac{1}{4}$ waveplate (crystalline or liquid crystal) and a liquid crystal retarder (retardance 0 and 180°).

The narrow band filter is the major design subject in NIRTF. Two choices considered in the design are Fabry-Perot Interferometer (FPI) and Lyot filter. Both of them are capable of providing high spectral resolution. The mechanism of FPI is the multiple-reflection

Table 2.1 Scientific Requirement of NIRTF

Parameters	Specs
Spectral coverage range	1.0~1.7μm
Spectral resolution	$\lambda/\Delta\lambda = 150,000$
FOV	1~3 $''$
Bandpass	0.1Å
Spatial resolution	<0.1$''$
Multi-operational mode	broad, medium, narrow (Table 2.2)
Filter aperture	Lyot filter ~ 36mm, FPI~150mm
Tunable over	1.0~1.7μm
High throughput	>20%
Stray light	$\leq 10^{-3}$

between two mirrors. The advantages of using FPI is that a very high spectral resolution can be achieved. However, narrower bandpass always comes with smaller free spectral range (FSR, the distance between two adjacent transmission peaks in spectral space). In order to select a specific bandpass of FPI, another filter is required. Multiple FPIs design is based on such an idea. There has been a recent development of a triple FPI system present in Gary et al. (2007). In the design of NIRTF, another approach is taken — FPI plus Lyot filter. In addition to the interference prefilter, a Lyot filter is designed as another prefilter for FPI. The advantage is that Lyot filter has simpler mechanical structure and easy to be calibrated.

The NIR imaging array had been a problem for years. Recently, there are breakthroughs in the development of NIR detectors. Presently, the NIR detector used at BBSO is a HgCdTe NIR CMOS designed by Rockwell, Imaging Science Division, and it has high quantum efficiency from 10000Å to 17000Å. The image format taken by this CMOS chip can be 1024×1024.

In the following sections, NIR FPI and Lyot filter will be discussed in detail. More information about polarization analyzers can be found in Spirock (2005). And the design of NIR imaging array is not included in the current presentation, and only relevant technique parameters of such imaging array will be included in the following discussion.

2.2.1 Near Infrared Fabry-Perot Etalon

In recent years, Fabry-Perot etalons have gained extensive interest in solar astronomy community due to the extraordinary performance that a FPI possesses: very high spectral resolution, high throughput, easy wavelength tuning, and large aperture. A typical Fabry-Perot etalon is shown in Figure 2.1.

The interference pattern of FPI is determined by:

$$2nd\cos\theta = m\lambda \tag{2.5}$$

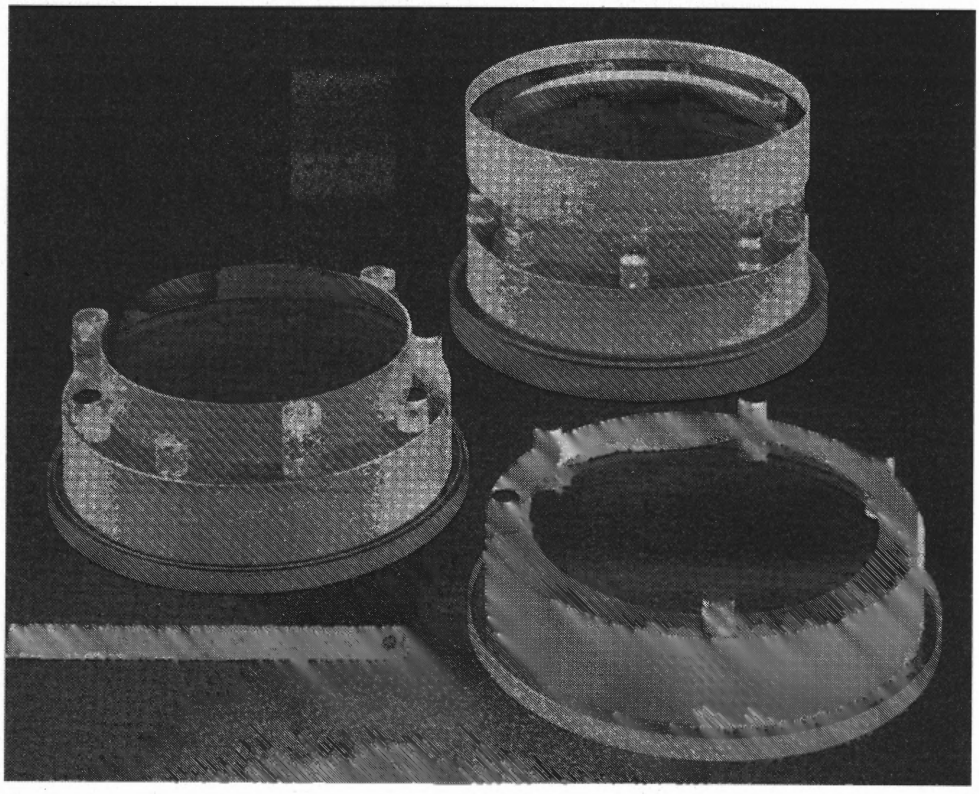

Figure 2.1 Fabry-Perot etalons. Upper-right is an assembled etalon; lower-left are two pieces from a single etalon.

where n is the refractive index of the medium between two flat mirrors; d is the distance between the mirrors; θ is the incident angle; m is the order of the interference fringe; λ is the wavelength. The transmission is:

$$\frac{I^t}{I^i} = \left(1 - \frac{A}{1-R}\right) \frac{1}{1 + \frac{4R}{(1-R)^2} \sin^2 \frac{\delta}{2}} \tag{2.6}$$

where, R is the reflectivity, A is absorption and scattering losses, $\delta = 4\pi nd\cos\theta/\lambda$ is the phase difference between successive beams, I^t stands for transmitted intensity, I^i is the intensity of the incident light. Free spectral range (FSR, the distance between the two

adjacent transmission peaks) is:

$$FSR = \frac{\lambda_0}{m} = \frac{\lambda_0^2}{2nd\cos\theta} \tag{2.7}$$

The width of the bandpass is measured with full width at half maximum (FWHM):

$$FWHM = \frac{2}{\pi\sqrt{\frac{4R}{(1-R)^2}}}\frac{\lambda_0}{m} = \frac{FSR}{\pi\sqrt{\frac{4R}{(1-R)^2}}} \tag{2.8}$$

And a special parameter — fineness — is defined as $F = \frac{FSR}{FWHM}$. An example of these properties of a FPI calculated for aperture 150mm and for $1'$ FOV is show in Table 2.3. According to the parameters in Table 2.3, the bandpass properties of the FPI are:

- At 15648.5Å, FSR = 5.5Å; FWHM=0.085Å.
- At10830Å, FSR=2.636Å; FWHM=0.04Å.

The transmission profile of the same FPI can be calculated from Equation 2.6 for these two spectral regions as shown in Figure 2.2.

2.2.2 Near Infrared Tunable Birefringent Filter

A birefringent filter depends on the interference of the polarized light transmitted through slabs of birefringent crystals in the direction perpendicular to the plane on which the optic axes sit. Let n_o and n_e be the ordinary and extraordinary indices of refraction. The difference between n_e and n_o, $\mu = n_e - n_o$ is known as the birefringent index of the crystal. There are two kinds of birefringent crystals according to the sign of μ, positive ($\mu > 0$ such as quartz) and negative ($\mu < 0$ such as calcite) crystals. Consider a plate of some birefringent crystal which is cut with its surface parallel to the optical axis of the crystal and perpendicular to the incident light ray. Its optical axis is 45° with respect to the polarization directions of the polarizers, placed between two parallel polarizers. The light ray in the crystal will

Table 2.2 Observation Modes of NIRTF

Observing Modes	Bandpass	FOV	Solar Lines (nm)	Scientific Interests
Blocking Filter	$20 \sim 40$Å	–	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	$\sim 120''$	FeI-1564.85, 1565.24, HeI-1083.0, etc. CN at J-band, OH at H-band.	Line core/wing images
Blocking Filter & Birefringent Filter & FPI	0.085 Å@ 1565.0 nm, 0.045 Å@ 1083.0 nm	$\sim 60''$	FeI-1572.36, 1565.29, 1564.85, 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.

Table 2.3 Specification of the NIR FPI for NIRTF of ATST

Parameters	Values
Clear Aperture	150 mm
Surface Quality	$\lambda/100$ at 633 nm
Wedge Angle	0 ± 1 fringe
Mirror Spacing	2226 μm (nominal)
Cavity Tuning Range	> 4.1 μm
Operational Temperature Range	0 °C to 50 °C
Nominal Finesse	>60 at 1523 nm
Coating Reflectivity	96%±1% from 1000 to 1700 nm

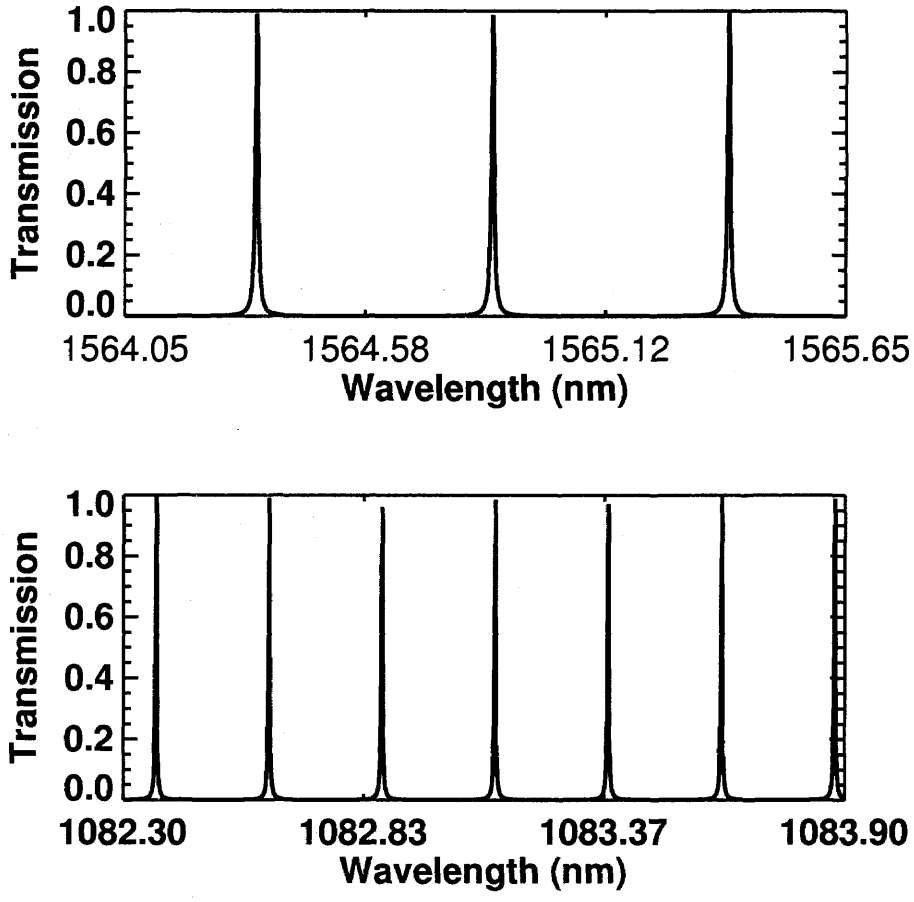

Figure 2.2 FPI of NIRTF for ATST, defined in Table 2.3.

be split into two rays, one is extraordinary light and the other is ordinary light. The fraction of transmitted light $\tau_{\parallel}$ will be a function of wavelength λ, and temperature, given by:

$$\tau_{\parallel} = \cos^2 \frac{\delta}{2} = \cos^2 \frac{\pi\mu(\lambda, T)}{\lambda} = \cos^2(\sigma\pi) \tag{2.9}$$

where, d is the thickness of the crystal, μ is the birefringent index of the crystal and λ is the wavelength of light. δ is known as phase difference between the e-ray and o-ray:

$$\delta = 2\pi \frac{(n_e - n_o)d}{\lambda} = 2\pi \frac{\mu d}{\lambda} = 2\pi\sigma \tag{2.10}$$

where $\sigma = \mu d/\lambda$ is called the retardation of a birefringent retarder. $\tau_{\parallel}$ will be maximized when σ is an integer and minimized when σ is half an integer. This polarizer $\leftrightarrow$ retarder $\leftrightarrow$ polarizer unit represents a basic building block of a birefringent filter.

Usually a birefringent filter is composed of several such units with the only difference being the thicknesses of the crystals. One setup is to set $d_0 = 2d_1 = 4d_2 = 8d_3 = \cdots$. The transmittance of the filter is therefore:

$$\tau_{\parallel} = \cos^2(\pi\sigma_0)\cos^2(\pi\sigma_1)\cos^2(\pi\sigma_2)\cos^2(\pi\sigma_3) \tag{2.11}$$

The bandpass of such a birefringent filter can be brought down to a fraction of an angstrom. The FWHM of the filter is determined by the thickest unit, and the FSR of the filter is determined by the thinnest unit of the filter.

2.2.3 Tuning Mechanism of a Birefringent Filter

In Equation 2.11, by varying the values of σ_i's, the transmission of the filter will be shifted. In order to change σ_i's, either the $\mu(\lambda, T)$ or the d need to be changed. This can be realized using two different approaches.

First, use a rotating half waveplate in each stage, as shown in Figure 2.3. 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.

Let α be the azimuthal angle of the rotating half waveplate. The transmission profile of a single unit is:

$$\tau_{\parallel} = \cos^2\left(\frac{\mu(\lambda,T)d}{\lambda}\pi + 2\alpha\right) = \cos^2(\sigma\pi + 2\alpha) \tag{2.12}$$

The rotating half waveplate can be set up on a high speed motorized stage. An example is the Newport RGV100 Series Motorized Rotation Stage, which has rotation speed of 720 °/s. This rotation speed results in a phase shifting rate of 8 waves/s, i.e., 125 ms/wave.

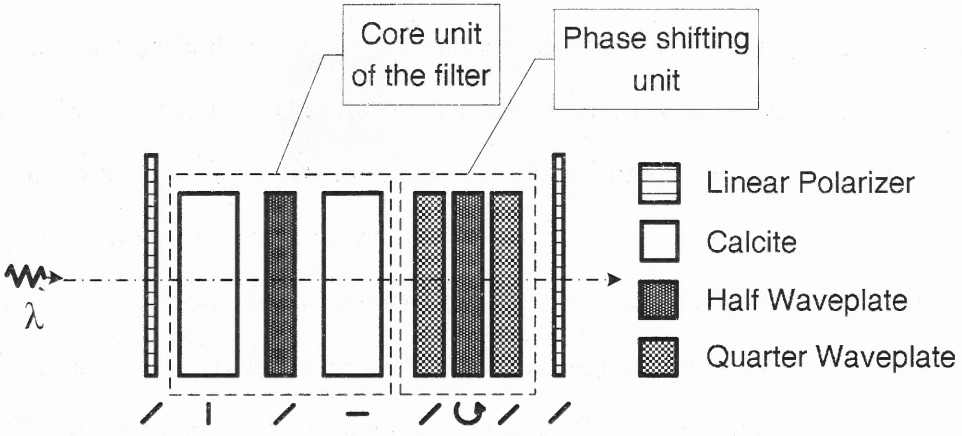

Figure 2.3 Concept design of a single stage birefringent filter.

There are several advantages of using a mechanically rotating half waveplate. 1) The change in α is linear; 2) The waveplate can be set in silicon oil which is usually applied in birefringent filter for temperature concerns; 3) The optical quality of the waveplate can be very high quality for imaging application. The disadvantages are: 1) The waveplate has to be achromatic; 2) The motorized stage increases the size of the birefringent filter; 3) The responding time is long (compared to the other tuning method below).

The second approach to tune the bandpass of a birefringent filter is to use liquid crystal (LC) variable retarder in place of the rotating half waveplate.

$$\tau_{\parallel} = \cos^2\left[\left(\frac{\mu(\lambda,T)d}{\lambda}+\Delta\right)\pi\right] = \cos^2(\sigma+\Delta)\pi \tag{2.13}$$

where, Δ is the phase change introduced by the LC variable retarder. The advantages of using LCVR are: 1) More compact design; 2) The response time is quick. The disadvantages are: 1) Nonlinear phase change introduces challenging calibration procedure; 2) Hard to integrate the LCVR with waveplates; 3) Non-uniformity due to stress within LCVR degrades

the image quality.

2.2.4 Design Parameters of the Birefringent Filter of NIRTF

In order to be compatible with FPI, the tuning range of birefringent filter is also from 1000 nm to 1700 nm. The challenges regarding the achromatism of the optical components concern waveplates and polarizers in the filter. Due to the fact that polarizers and calcites cannot be made in large size, the clear aperture for the filter is chosen to be 36.0 mm. The design specifications of the birefringent filter proposed for NIRTF of ATST are as shown in Table 2.4.

In order to match the FSR of FPI in Table 2.3, which is 5.5 Å at 15648.5Å, the FWHM of the birefringent filter has to be at least 5.5 Å. As mentioned before, the FSR of the birefringent filter is determined by the thickness of the thinnest module; the FWHM is determined by the thickest module.

Note that in Table 2.5, the FWHM of the birefringent filter is not the FWHM of the thinnest module, instead, in combination with the bandpass profiles of other filters, it is down scaled by a factor of 0.883~0.887. FSR is the same as the FSR of the thickest module. The value of $\mu(\lambda, T)$ for calcite was investigated in Wang et al. (2001):

- $\mu(15650, 35°C) = -0.171667$
- $\mu(10830, 35°C) = -0.175250$

Table 2.4 Design Parameters for Lyot Filter in NIRTF

Parameters	Values
Aperture:	ϕ36.0 mm
FWHM:	~2.4Å
Working Temperature:	35±0.05 °C
FSR:	$\geq$ 40.0Å
Tunable Range:	1000 nm to 1700 nm
Transmission:	as high as possible

Due to this temperature sensitivity, birefringent filter is usually housed in a box enclosed by a temperature controller. The FSR of the birefringent filter requires the bandpass of the interference filter to be as follows: For 15650Å, the off-band transmission should be 10^{-4} outside of the passband 15625Å to 15675Å — 15650±15Å; For 10830Å, 10^{-4} outside the band from 10815Å to 10845Å — 10830Å±10Å.

2.2.5 Wide-Field Configuration in Birefringent Filter

Generally, a birefringent filter works better with collimated beams better. The limitation is due to the fact that any non-zero incident angle can introduce extra phase changes in the filter, and effect the designed interference pattern. In order to reduce the dependence on the incident angle, a modified setup of a birefringent filter unit is usually applied due to Evans (1949). In this setup, the calcite plate is split into two identical plates, with the insertion of a half waveplate between them.

Consider a single birefringent crystal cut in a way such that its fast axis (optic axis) lies on the surface. The incident light polarized in a plane of 45° with respect to the fast axis enters the crystal in the direction (ϕ, θ), where ϕ is the angle of incidence angle and θ is the azimuth of the polarization plane measured from the fast axis (Figure 2.4). The light exits from the other surface of the crystal in the direction (ϕ, θ) in two polarized components with polarization closely parallel to the fast axis and slow axis respectively. The retardance between these two polarizations, δ, can be determined by (ϕ, θ) and R_0, where R_0 is the

Table 2.5 Design of The Birefringent Filter of NIRTF of ATST

Module	Thickness (mm)	FWHM(Å)	FWHM(Å)	FSR(Å)	FSR(Å)	Retardation	Retardation	Shape/Size (Octagon/mm)
		15648.5Å	10830.3Å	15648.5Å	10830.3Å	15648.5Å	10830.3Å	
0	12.560×2	2.751	1.332	5.502	2.664	2844.0	4046.8	37.00×37.00
1	6.280×2	5.502	2.664	11.004	5.328	1422.0	2032.4	37.00×37.00
2	3.140×2	11.004	5.328	22.008	10.656	711.0	1016.2	37.00×37.00
3	1.570×2	22.008	10.656	44.016	21.312	355.5	508.1	37.00×37.00
Final	–	2.43	1.180	44.016	21.312	–	–	–

retardance when the incident angle is zero. This can be shown in the following equation:

$$\delta = \delta_o \left[1 + \phi^2 k \left(\frac{\cos^2\theta}{n_o} - \frac{\sin^2\theta}{n_e} \right) \right] \tag{2.14}$$

where, $k = -\frac{1}{2n_o}$ for calcite. For example,

$$\begin{aligned} \phi_0^2 &= \left(\frac{\delta(\phi_0, \theta = 0)}{\delta_0} - 1 \right) \frac{n_o}{k} = \left(1 - \frac{\delta(\phi_0, \theta = 0)}{\delta_0} \right) 2n_o^2 \\ \phi_{\pi/2}^2 &= \left(1 - \frac{\delta(\phi_0, \theta = \pi/2)}{\delta_0} \right) \frac{n_e}{k} = \left(1 - \frac{\delta(\phi_0, \theta = \pi/2)}{\delta_0} \right) 2n_o n_e \end{aligned}$$

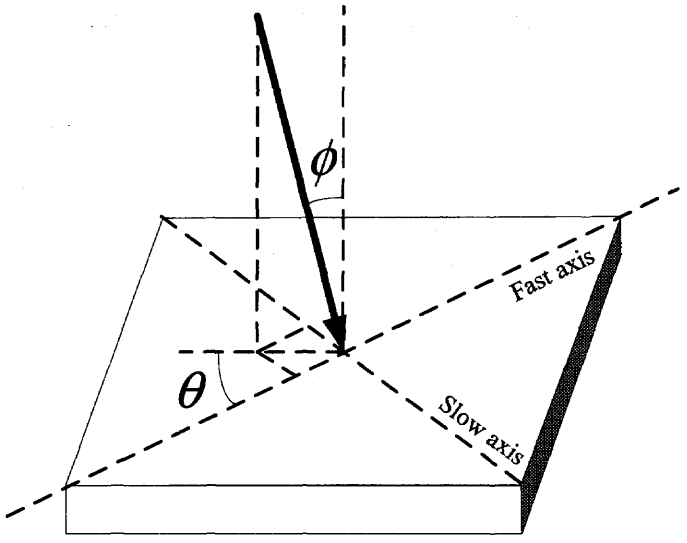

Figure 2.4 Inclined incidence light on a birefringent crystal.

In the wide field configuration, the optic axes of the two half calcites are crossed and the optic axis of the inserted half waveplate is 45°with respect to both of the axes of the calcites. Therefore, given the incident light at (ϕ, θ), it will incident on the second half

at angles $(\phi, \theta + \frac{\pi}{2})$. Let R be the retardation of the overall calcite crystal,

$$\begin{aligned} \delta &= \frac{1}{2}\left[\delta(\phi,\theta)+R(\phi,\theta+\frac{\pi}{2})\right] \quad (2.15) \\ &= \frac{1}{2}\delta_0\left[1+\phi^2 k\left(\frac{\cos^2\theta}{n_o}-\frac{\sin^2\theta}{n_e}\right)\right]+\frac{1}{2}\delta_0\left[1+\phi^2 k\left(\frac{\sin^2\theta}{n_o}-\frac{\cos^2\theta}{n_e}\right)\right] \\ &= \delta_0\left[1+\phi^2 k\left(\frac{1}{n_o}-\frac{1}{n_e}\right)\right] \quad (2.16) \end{aligned}$$

Note that for calcite, $n_e < n_o$, and $k = -\frac{1}{2n_o}$ is negative, so $R > R_0$. Further, it can be derived that

$$\phi^2 = \left(\frac{\delta}{\delta_0}-1\right)\frac{n_o n_e}{k}\frac{2}{n_e-n_o} = \left(\frac{\delta-\delta_0}{\delta_0}\right)\frac{4n_o^2 n_e}{n_o-n_e} \quad (2.17)$$

Therefore, for constant retardation δ, the incident angel ϕ for wide field configuration is $\frac{2n_e}{n_o-n_e}$ larger for $\theta = 0$ and $\frac{2n_o}{n_o-n_e}$ larger for $\theta = \pi/2$. Specifically, given $n_o = 1.63337$ and $n_e = 1.47764$ for calcite at 1564 nm (Ghosh 1999), $\frac{2n_e}{n_o-n_e} \approx 19$ in $\theta = 0$ case, and $\frac{2n_o}{n_o-n_e} \approx 21$ in $\theta = \pi/2$ case. Therefore the curvature of the parabola is flattened in the wide field configuration, and the ϕ dependence is reduced (θ dependence is eliminated!).

For the near infrared birefringent filter designed for NIRTF/ATST, there are four modules with retardation 2844, 1422, 711, and 356 (Table 2.5). Let $\Delta\delta = \mid \delta - \delta_0 \mid = 0.01$, which corresponds to 0.05Åat 15648Å(0.027Åat 10830Å) shift of the bandpass, the corresponding incident angles to each module can be tabulated as shown in Table 2.6.

Table 2.6 FOV Designs of Lyot Filter Modules

Wavelength (nm)	Module 0	Module 1	Module 2	Module 3
1042	1°01′	1°27′	2°03′	2°54′
1097	1°00′	1°26′	2°02′	2°52′
1159	1°01′	1°26′	2°02′	2°53′
1220	1°01′	1°26′	2°02′	2°53′
1273	1°00′	1°25′	2°00′	2°52′
1307	1°00′	1°25′	2°01′	2°51′
1497	1°00′	1°25′	2°01′	2°51′
1541	1°00′	1°25′	2°01′	2°51′
1609	1°00′	1°25′	2°01′	2°51′
1682	1°00′	1°25′	2°01′	2°51′

2.2.6 F-ratio of the Birefringent Filter of NIRTF

The optical setup can be telecentric as well, see Figure 2.5. In this case, the focal plan of the incoming light lies within the filter. The incident angles are not small anymore to the filter, which have two effects: 1) side bands appear; 2) clear aperture reduces.

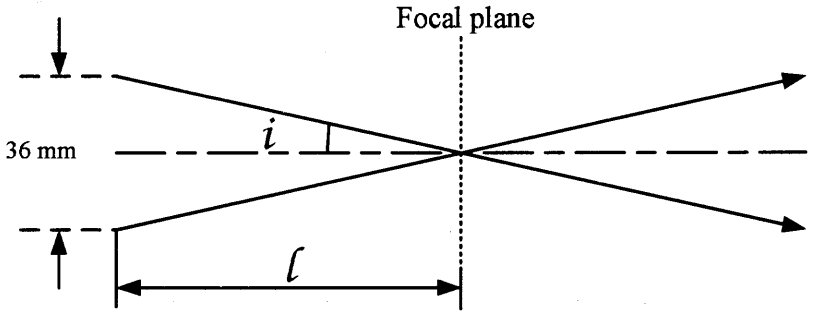

Figure 2.5 Inclined incidence light on a birefringent crystal.

Assume the clear aperture of the filter is $\phi 36$ mm, the allowed maximum length of the filter is $2l$, then

$$\tan i = \frac{1.8\text{cm}}{l} \tag{2.18}$$

The f/ratio is then:

$$\text{f/ratio} = \frac{1}{3.6\text{cm}} \sim \frac{180}{\pi \cdot 2\text{i}} \tag{2.19}$$

In practice, the length of the filter is much less than 100 cm, therefore the beam speed can be set to f/29 and up.

2.2.7 Near Infrared Polarizer of NIRTF

Polarizers are the major sources responsible for large amount of absorption, especially in near IR spectral range. The traditional thin-film polarizer made by 3M (formerly Polaroid) has satisfying performance in visible range, but very poor in near IR range — low transmission and low extinction ratio.

Recently, a newly developed nano-technology made it possible to manufacture high

performance glass polarizer. For examples, Polarcor of Corning (NY/USA); colorPol of Codixx (Barleben/Germany); SubWave polarizer of Nanoopto (NJ/USA). The major threshold at this stage is the dimension of these glass polarizers can only be on the order of 1 to 2 inches square. Currently, the 1500 BC5 colorPol of Codixx holds transmission from 85% to 89% from 1000 nm to 1700 nm without anti-reflection coating. In the application of the birefringent filter, the polarizer with be contacted with silicon oil or grace in order to match the refractive indices of optical components to reduce surface reflection. The reflection of the polarizer on each interface is 4% for each surface. Therefore the total transmission is 93%$\sim$97%. At the same time, the extinction ration (ER) is $> 10000:1$ from 1000 nm to 1700 nm.

The NIRTF birefringent filter holds four modules — five polarizers in total. The transmission of the filter can be accounted as follows: 1) For polarized light and ignore other losses, $\tau = (93\%)^5 \approx 69.57\%$ for 93% polarizer, and $\tau = (97\%)^5 \approx 85.87\%$ for 97% polarizer; 2) For unpolarized light, the formerly calculated τ need to be reduced further by 50%.

2.3 IRIM System

The Near Infrared Imaging Magnetograph (IRIM) is a system built at BBSO in 2004. The system has very similar structure as NIRTF. Only the design goal of IRIM was to look at the sun at the close vicinity of the wavelength 1565 nm, i.e., the achromatic requirements in IRIM is not so stringent as in NIRTF. The setup of the IRIM system in BBSO coudé room is shown in Figure 2.6.

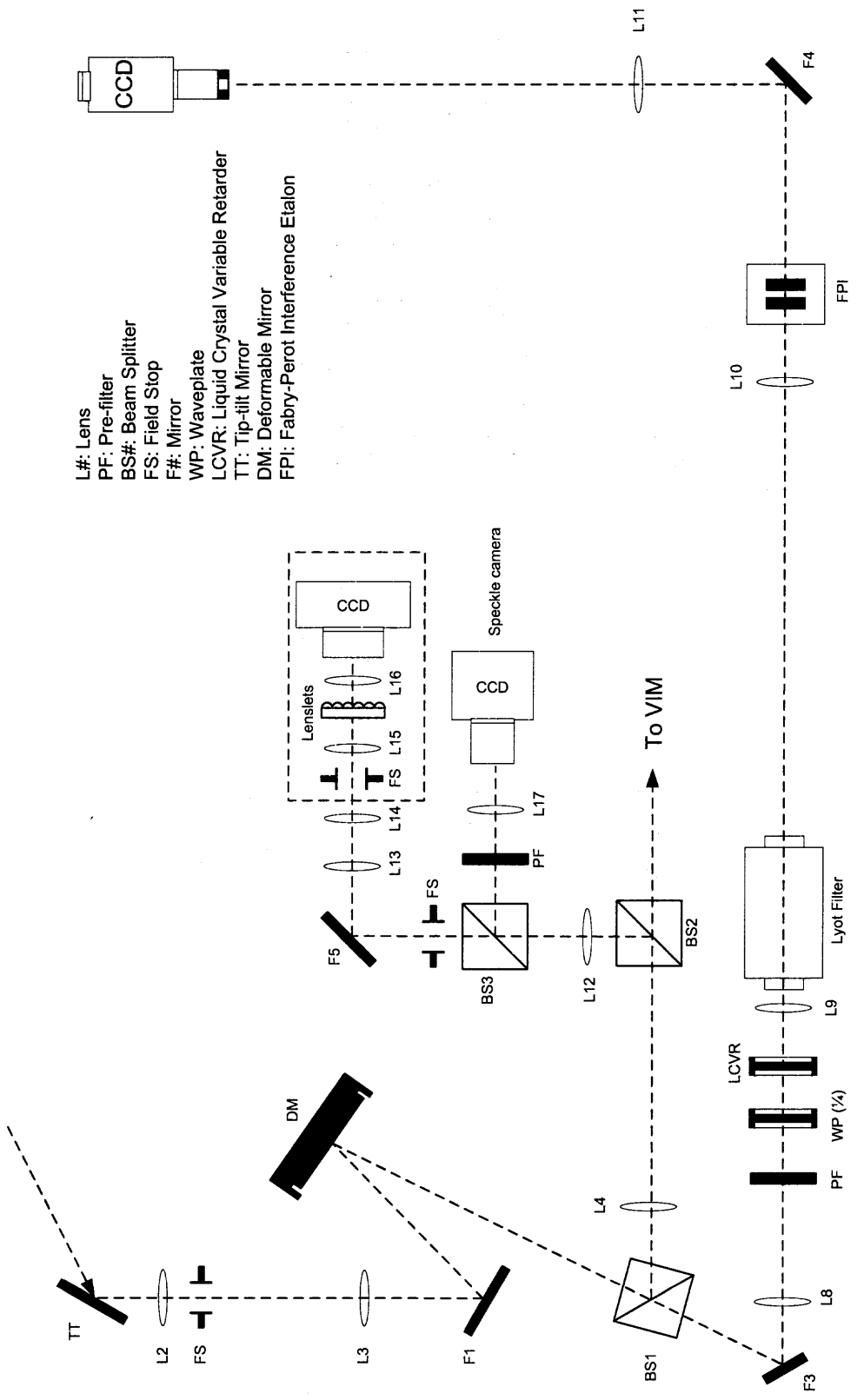

Figure 2.6 IRIM system setup in the Coudé room at BBSO (2004).

2.3.1 Bandpass Profile of Lyot Filter - Laser Testing

The filter was designed by Wang et al. (2001) and manufactured by Cambridge Research & Instrumentation, Inc. (CRi). It is a four-module LCVR tuning Lyot filter. Each module is installed on a cage, and the cages are then linked mechanically, see Figure 2.7. Therefore, it was possible to dissemble the filter in an optical lab with temperature controller. This was done at CRi's optical lab. A tunable laser was used to generate a continuously varying pseudo-monochromatic light for Lyot filter testing. The testing result is shown in Figure 2.8. The central wavelengths of the testing are 1564.87 nm (top) and 1565.28 nm (bottom). It was a point-testing, i.e., the diameter of the laser beam is small.

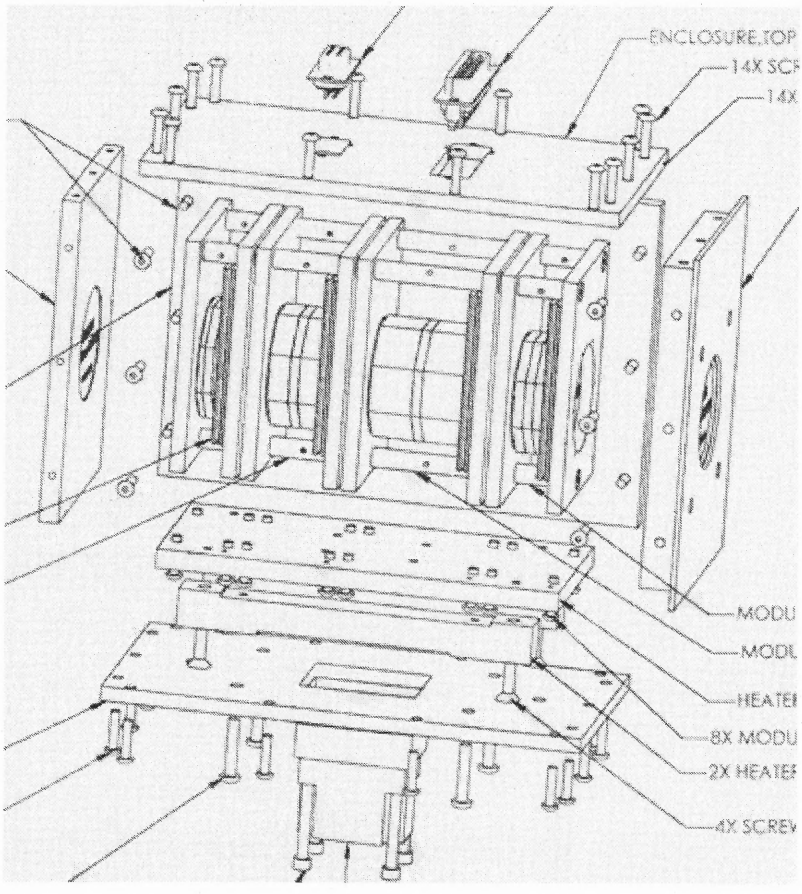

Figure 2.7 Mechanical drawings of the Lyot filter used in IRIM.

At CRi lab, the properties of LCVRs inside of the Lyot filter were also tested, see Figure 2.9, in which the y-axis unit is in nm. To convert into radians, simply divide the

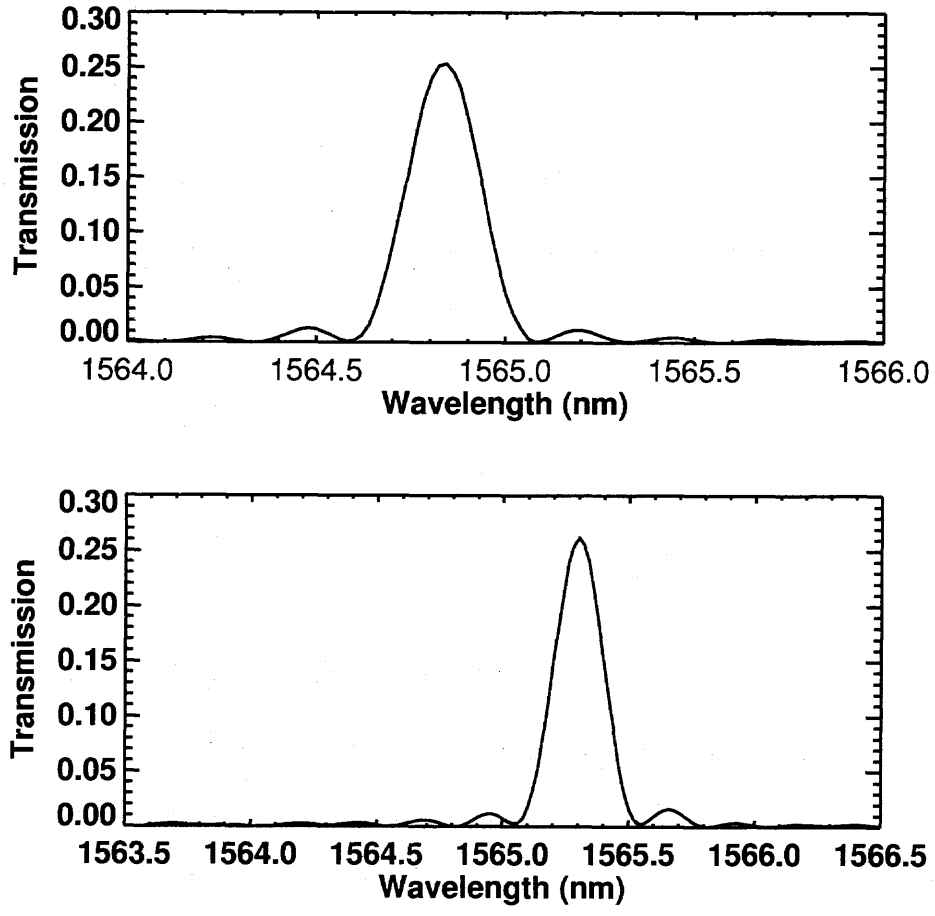

Figure 2.8 Laser testing of IRIM Lyot filter.

y-axis by the working wavelength.

2.3.2 Bandpass Profile of Lyot Filter - Spectrograph Testing

After the full assemble at CRi, the filter was further wrapped by a temperature controller. The need for such a temperature controller is that the birefringence of the crystals are functions of ambient temperature: $\mu = \mu(T, \lambda)$. Also, the LCVRs used for tuning purpose are also sensitive to the ambient temperature. Therefore, the calibration of the filter can only be done in a single unit after the temperature controller is installed.

The calibration results are shown in Figure 2.10 (The overall filter transmission is

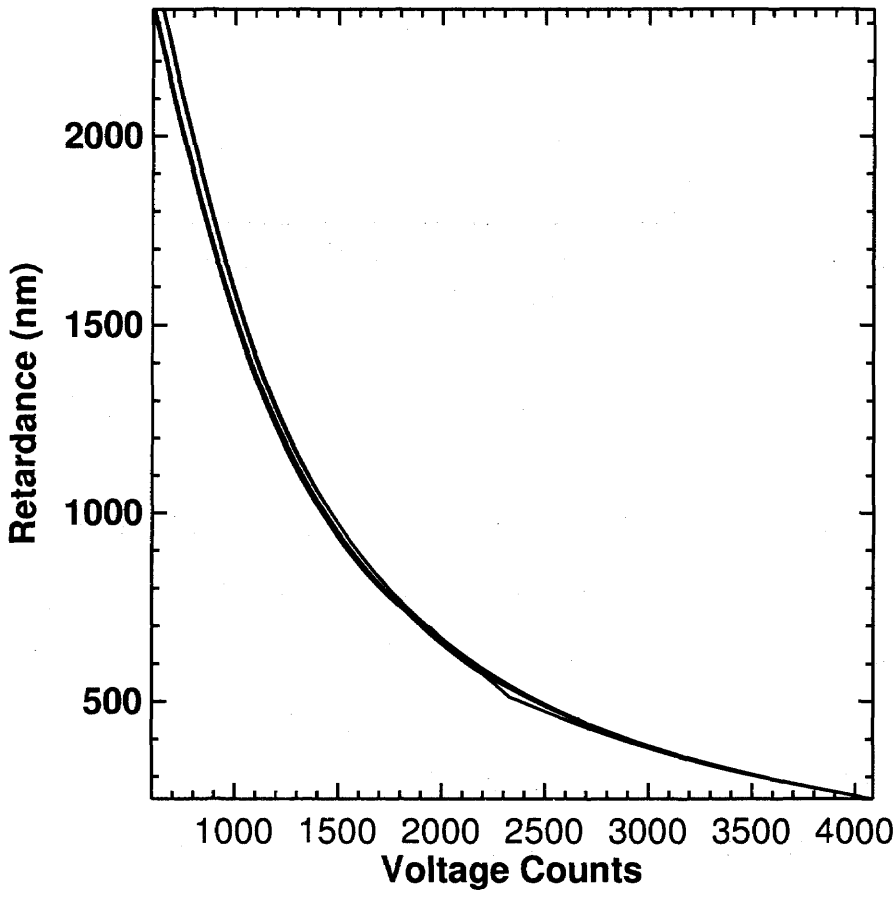

Figure 2.9 Testing of the LCVRs used in Lyot filter.

recoded as a function of the driving voltage (y-axis) of the thickest module and wavelength (x-axis). The voltage count from 0 to 500 were defected and do not create any response in LCVR. Ambient temperature is 32 °C.) and Figure 2.11 (The spectral lines shown in the figure are 15648.5Åand 15652.9Å. Each of the two images on the left is a pair of the spectrum with/without the Lyot filter.). Compared with the Figure 2.8, the first order maximum in the spectrograph tests are much higher than in the laser tests. Note that the spectrograph was done at CRi in April/2004, and the laser test was done at NSO/SP in December/2004. Also, in laser tests, only a few samples on the filter aperture were chosen. In spectrograph test, however, the intensity profile is an area average over the aperture. The

non-uniformity of the bandpass over the aperture can introduce this disparity. In fact, the source of scattering light and higher sideband of the Lyot filter can be identified using ring like experiment, i.e., every time, only let a ring of the light pass through the aperture.

2.3.3 Fabry-Perot Etalon of IRIM

The FPI used in IRIM is a 70 mm ET70FS-1041 series manufactured by Queensgate Instruments, Ltd. operated by a controller CS-100 to adjust the mirrors. The properties of the etalon are listed in Table 2.7.

2.3.4 Polarization Analyzer of IRIM

The analyzer of IRIM is a traditional composition of a $\lambda/4$ waveplate, a LCVR, and a linear polarizer. The combination serves the purpose of measuring longitudinal field only. The optic axis of the fixed $\lambda/4$ waveplate is parallel to the active direction of LCVR, and 45°with respect to the active direction of the entrance polarizer of Lyot filter. In the Zeeman triplet, the two circularly polarizer σ-components are converted to two orthogonal linear polarizations by the fixed $\lambda/4$ waveplate. Presumably there is no photon loss in this conversion. In order to record the photon flux in the two components, the LCVR is switched between two states: 0 and $\lambda/2$. In doing so, Stokes $I+V$ and $I-V$ is selected. The linear component π of Zeeman triplet will be blocked by the linear polarizer in the analyzer.

2.3.5 Magnetogram of IRIM

A sample magnetogram obtained on July 1st, 2005 (BBSO) is shown in Figure 2.12, in which, left panel is IRIM high resolution magnetogram (UT 16:02); the right panel is MDI (UT 16:00). The correlation tracking (CT) system and high-order adaptive optics (AO) (Ren et al. 2003; Didkovsky et al. 2003; Rimmele 2000; Rimmele et al. 2003) are applied

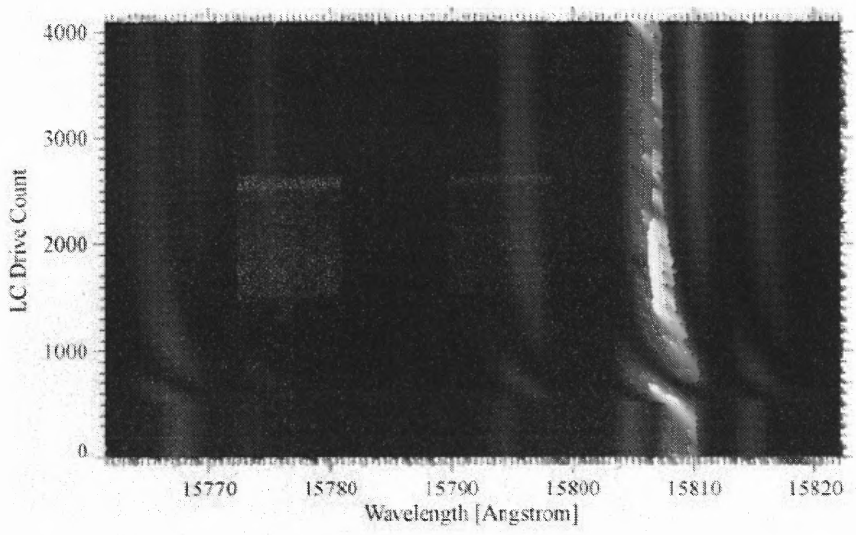

Figure 2.10 Spectrograph testing of the Lyot filter of IRIM - Liquid Crystal Tuning.

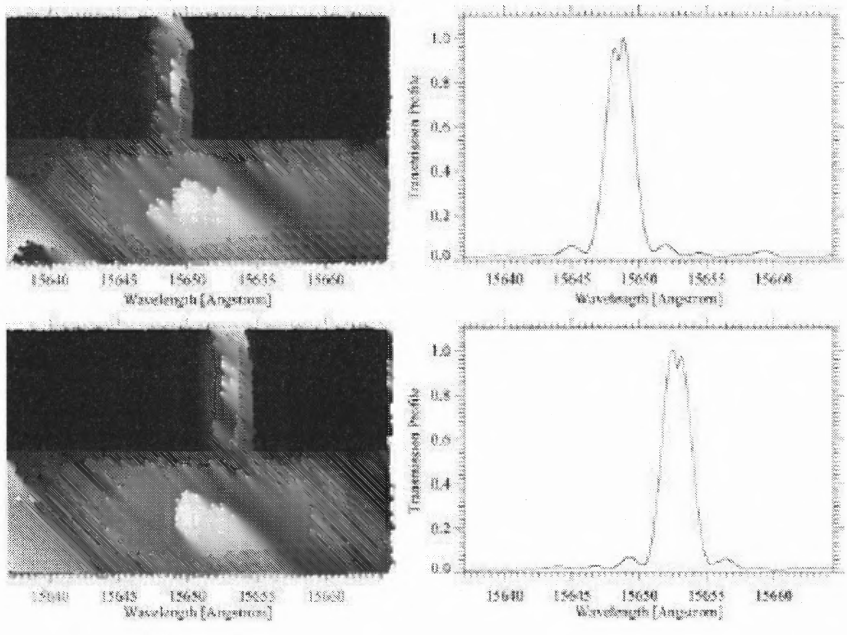

Figure 2.11 Spectrograph testing of the Lyot filter of IRIM.

in the observation. The CT system holds a 45 mm tip-tilt mirror, operated at 3.3 kHz, and tilt range of 24″. The AO system consists of a 77 mm deformable mirror (DM) operated by 97 actuators and a Shack-Hartmann wave-front sensor embedded with 76 subapertures. These systems enable a pleasant diffraction-limit observation at BBSO.

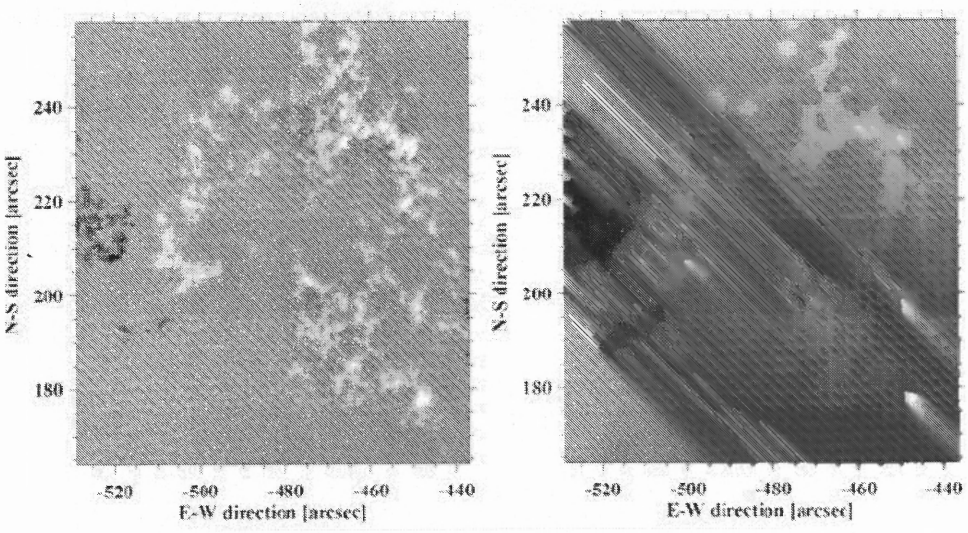

Figure 2.12 Magnetogram of NOAA AR 10781 obtained (Fe I 1564.85 nm).

Table 2.7 Fabry-Perot Etalon of IRIM (Cao et al. 2006)

Parameters	Specs
Diameter	70 mm
Nominal spacing	2226 μm
Scanning range	4.1 μm
Reflection	95.3%
Absorption	0.27%
Peak transmission	88.8%
FSR	0.55 nm
FWHM	8.8 pm
Finesse	62.8
Flatness	$\lambda/127$
Roughness	$\lambda/763$
Step size	0.29 pm/step

CHAPTER 3

DESIGN OF ACHROMATIC WAVEPLATES

The simplest form of an electromagnetic wave is a plane wave represented as follows:

$$\mathbf{E}(\mathbf{x},t) = \mathbf{E}_0 \exp(i\mathbf{k}\cdot\mathbf{x} - i\omega t) \tag{3.1}$$

where, $\mathbf{E}_0$ is the vector amplitude of the electric vector $\mathbf{E}(\mathbf{x},t)$ of the light. The $\mathbf{E}_0$ corresponds to the polarization of the light; $\mathbf{k}$ is related to the electromagnetic energy; and ω is related to the color of the light. The phase velocity $v_p = \frac{\omega}{k}$ describes the transportation the wave pattern; and the group velocity $v_g = \frac{\partial\omega}{\partial k}$ describes the energy transportation (where k is the module of the wave vector $\mathbf{k}$). The $\omega - k$ diagram is thus an important curve of medium properties, which depicts what kind of $\omega - k$ relation is supported by the specific medium. According to Fourier theorem, any form (square integrable) of the solutions of the wave equation can be represented by the summation of Fourier components — plane waves. Achromatic waveplate is such an optical component that can produce a predefined phase retardance between two orthogonal components (polarized components) of the amplitude vector of the light regardless of the wavelength of the light.

3.1 E&M Wave in Anisotropic Medium

The two entries of Maxwell equations of importance here are (Born & Wolf 1999):

$$\nabla\times\mathbf{E} + \frac{1}{c}\frac{\partial\mathbf{B}}{\partial t} = 0 \tag{3.2}$$

$$\nabla\times\mathbf{H} - \frac{1}{c}\frac{\partial\mathbf{D}}{\partial t} = \frac{4\pi\mathbf{j}}{c} \tag{3.3}$$

Given the medium is not magnetic active (μ =1 in Gaussian Unit), the medium properties is uniquely described by its dielectric tensor ε_{ik}. Therefore, $\mathbf{D} = \varepsilon\mathbf{E}$ and $\mathbf{B} = \mathbf{H}$. Source

free says $\mathbf{j} = 0$. According to the monochromatic plane wave representation, operators $\nabla\times \rightarrow i\mathbf{k}\times$, and $\frac{\partial}{\partial t} \rightarrow -i\omega$. Therefore,

$$\mathbf{k}\times\mathbf{E} = \frac{\omega}{c}\mathbf{B} \tag{3.4}$$

$$\mathbf{k}\times\mathbf{B} = -\frac{\omega}{c}\mathbf{D} \tag{3.5}$$

Let $\mathbf{k} = \frac{\omega}{c}\mathbf{n}$, then

$$\mathbf{n}\times(\mathbf{E}\times\mathbf{n}) = \mathbf{D} \tag{3.6}$$

The magnitude of the vector $\mathbf{n}$ is still termed as refractive index, although a simple refraction law does not exist for anisotropic medium. Expand this equation and consider $\mathbf{D} = \varepsilon\mathbf{E}$,

$$n^2\delta_{ik}E_k - n_kE_kn_i = \varepsilon_{ik}E_k \tag{3.7}$$

The condition to have non-trivial solution of $\mathbf{E}$ requires the secular equation stands:

$$\det \mid n^2\delta_{ik} - n_in_k - \varepsilon_{ik} \mid = 0 \tag{3.8}$$

The secular equation is of the form of an ellipsoid in general. Figure 3.1 shows a wave normal surface of an uniaxial medium (Born & Wolf 1999). Wave normal surface is derived from the by utilizing the relation $\mathbf{n}\cdot\mathbf{s} = 1$. There is no physical motivation to favor a normal surface, but for the sake of mathematical clarity only. Normal surface is a surface of refractive index. $\mathbf{s}$ surface is a surface of phase velocity. The correlation relation between two surfaces are: $\mathbf{n} \leftrightarrow \mathbf{s}$, $\mathbf{E} \leftrightarrow \mathbf{D}$ and $\varepsilon_{ij} \leftrightarrow \varepsilon_{ij}^{-1}$. The plane Γ is formed by vectors $\{\mathbf{s},\mathbf{z}\}$, $\mathbf{D}'' \parallel \Gamma$ and $\mathbf{D}' \perp \Gamma$ (electrical induction vector $\mathbf{D} = \varepsilon\mathbf{E}$). The projection point p is element of plane Ξ. The normal of the circular cross section Z is the optical axis $\mathbf{z}$. Usually, the light passes along the x- or y-axis in the case of a birefringent filter. For a positive birefringent material such as Quartz $v_o > v_e$ (or $n_o < n_e$).

In a summary, the propagation of electromagnetic waves inside anisotropic medium

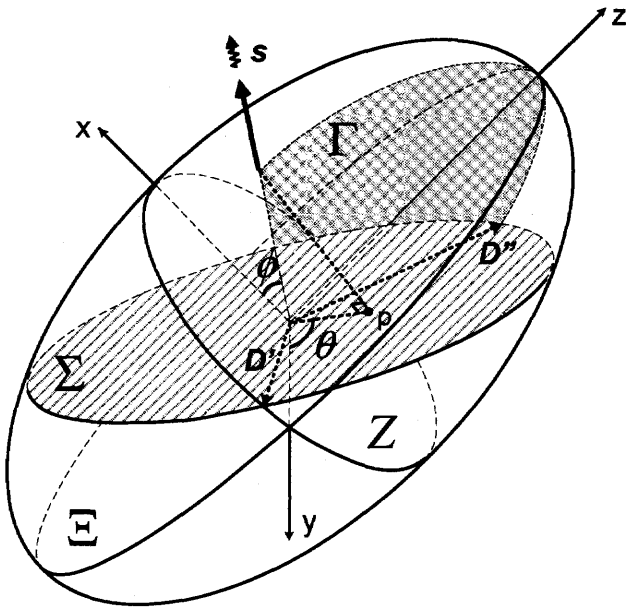

Figure 3.1 Ellipsoid of wave normals.

can take maximum three modes. Each mode has a corresponding phase velocity $v_p = \omega/k = c/n = c \cdot s$, which depends on the medium and the frequency of the wave itself. Note in Figure 3.1, the tip of the **s** vector can not leave the surface of the ellipsoid.

3.2 Polarization Optics

An optical system is usually considered as a linear system if only certain optical components are involved, such as retarders (waveplates), rotators, partial polarizers, etc.. There are many mathematical formalisms being developed to describe such linear system. The monograph by Shurcliff (1966) gives a thorough discussion on the subjects of Jones and Mueller calculus. The connection between these two popular formalisms is called the coherence matrix, which is a subject of statistical optics. In this section, these formalisms will be reviewed briefly for the convenience of later discussion.

3.2.1 Jones Calculus

Jones calculus was developed in a series of papers (Jones 1941a,b, 1942, 1947) and the basics principle of the calculus is that certain optical components can be though of as linear operators operating on polarized light. In cases where only plane waves are involved, the transformation can be represented as 2×2 matrices. These matrices are named Jones matrices after the name of the inventor. The transformation of the E-vectors of the light can be show as follows:

$$E_{x'} = E_x \exp\left[-i\frac{2\pi d}{\lambda}(n_x - i\varepsilon_x)\right] \tag{3.9}$$

$$E_{y'} = E_y \exp\left[-i\frac{2\pi d}{\lambda}(n_y - i\varepsilon_y)\right] \tag{3.10}$$

where, $n_x, n_y, \varepsilon_x, \varepsilon_y$ are assumed to be the principle refractive indices and extinction coefficients, respectively; d is the geometric length of the light path; λ is wavelength as usual. Written in matrix form: $\mathbf{E}' = \mathbf{J} \cdot \mathbf{E}$. Major conclusions developed in a series of papers are as follows:

- Theorem 1 —The transformation matrices in an optical system including only retarders and rotators are all unitary.
- Corollary 1.1 —Such a system conserves the intensity of the light, i.e., the energy.
- Theorem 2 —Given certain wavelength of light, an optical system including only retarders and rotators is equivalent to a system composed of one retarder and one rotator.
- Theorem 3 —Given certain wavelength of light, an optical system including only partial polarizers and rotators is equivalent to a system composed of one partial polarizer and one rotator.

- Theorem 4 —Given certain wavelength of light, an optical system including only retarders, rotators, and partial polarizers, is equivalent to a system composed of four optics: two retarders, one partial polarizer, and one rotator. In many cases, the rotator is even not necessary.

A phase retarder can be written as the Jones matrix

$$\begin{bmatrix} e^{i\delta/2} & 0 \\ 0 & e^{-i\delta/2} \end{bmatrix} , \tag{3.11}$$

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

$$\begin{bmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{bmatrix} , \tag{3.12}$$

where θ is the active direction of the polarizer.

The Jones matrix for a retarder with an arbitrary azimuthal angle θ is:

$$\mathbf{J} = \mathbf{R}(-\theta)\cdot\begin{bmatrix} e^{i\delta} & 0 \\ 0 & e^{-i\delta} \end{bmatrix}\cdot\mathbf{R}(\theta) , \tag{3.13}$$

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} . \tag{3.14}$$

3.2.2 Mueller Matrix

A Mueller matrix $\mathbf{M}$ is a 4×4 real matrix, which represents an optical instrument. In fact, a Mueller matrix can also describe the optical properties of a medium, such as the atmosphere of the Sun. Any Jones matrix has a corresponding Mueller matrix. The statement is not

true vice verse. Given a Jones matrix **J**, its corresponding Mueller matrix form is:

$$\mathbf{M} = \left[\mathbf{T}\cdot(\mathbf{J}\otimes\mathbf{J}^*)\cdot\mathbf{T}^{-1}\right]^T \quad , \tag{3.15}$$

where

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 \\ 0 & -i & i & 0 \end{bmatrix} \quad . \tag{3.16}$$

where "$\otimes$" denotes the Kronecker product, "$*$" denotes the operation of complex conjugate, and superscription "T"denotes matrix transpose.

Furthermore, if the vector representation of the light known to be the Jones vector **L** (a complex column vector), then the Stokes vector **S** for the same light in Mueller algebra is

$$\mathbf{S} = \mathbf{T}(\mathbf{L}\otimes\mathbf{L}^*) \quad . \tag{3.17}$$

The applications of Mueller matrix extends to not only the instrumental solar physics, but some theoretical sections such as radiative transfer equations. Mueller calculus gains these popularity mainly from its corresponding vector — Stokes vector. The four components of the Stokes vector are all observable and real. Usually, the vector is written as $\mathbf{S} = \{I, Q, U, V\}^T$. This vector will be discussed in more details in later chapters where the methodology of the magnetic field measurements is discussed.

3.3 *n*-Layer AWP

3.3.1 Methods

There are basically two approaches to the optical design of achromatic waveplates: combinatorial and structural waveplates. A combinatorial waveplate (see Figure 3.2) 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 wavelength dependence of the retardance. This is the method used in the current study and 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, for examples, Flanders (1983); Kikuta et al. (1997). This approach relies on semiconductor manufacturing technology, which is beyond the scope of this presentation.

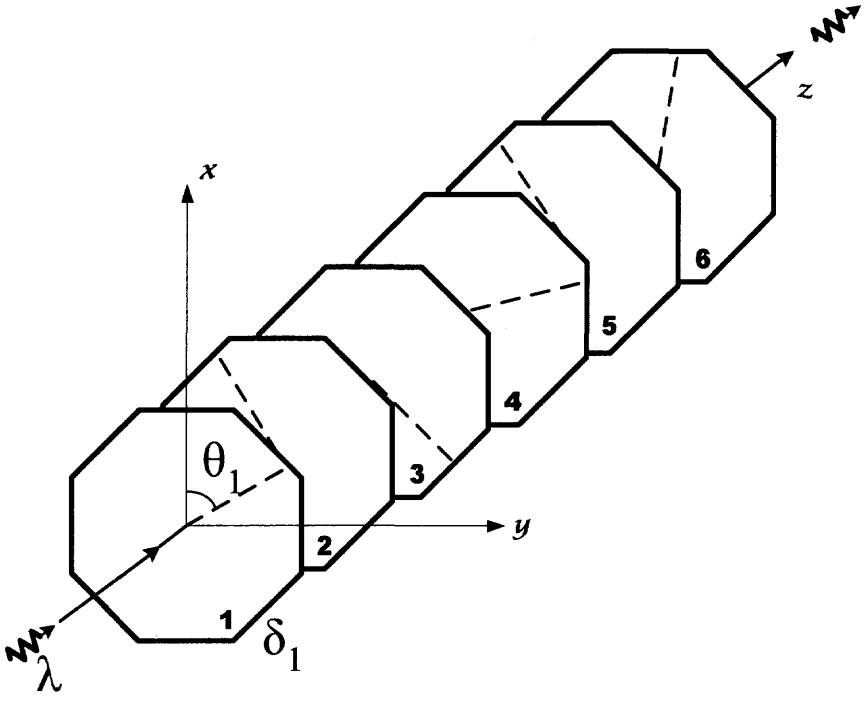

Figure 3.2 Diagram of a six-layer waveplate.

The idea of using three layer of birefringent material to achieve achromatic wave-

plate was initiated by Pancharatnam (1947). The methods of developing multilayer achromatic waveplate were further developed by various authors. For example, an analytical design of three-layer achromatic waveplates can be found in Title (1975), in which the author specified retardance of 115.5°, 180° and 115.5° and azimuth of 0°, 70.6° and 0° for a three-layer $\lambda/4$ waveplate. 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, Beckers 1971, 1972; Hariharan 1995; Guimond & Elmore 2004) or prisms (Filinski & Skettrup 1984). One issue in these designs is that optic-axes of the assembled waveplates also vary with wavelength.

3.3.2 Optimization Problem of *n*-layer Achromatic Waveplates

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.

Multilayer (more than three layers) waveplate discussed in the current paper is a solution to this issue. Our definition of an achromatic waveplate is that a multilayer structure which has a constant Jones matrix within a certain wavelength range. The definition implies the equivalency between a Jones matrix and a waveplate. In Figure 3.2, $\mathbf{J}_{\text{awp}} = \mathbf{J}_1 \cdot \mathbf{J}_2 \cdot \mathbf{J}_3 \cdot \ldots \cdot \mathbf{J}_n$ stands for the total Jones matrix of the multilayer waveplate and $\mathbf{J}_{\text{perfect}}$ is 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 $k = 1, \ldots, n$, which are the azimuth and retardance of each layer, respectively. The design criterium of an n-layer achromatic waveplate is to minimize the following merit function

$$E(\theta_k, \delta_k) = \sum_{\lambda,i,j} \Delta J_{ij} \cdot \Delta J_{ij}^* |_\lambda \tag{3.18}$$

$$\text{with} \quad \delta_k = 2\pi\mu d_k/\lambda \quad . \tag{3.19}$$

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 linear relation between δ and λ^{-1} is only valid for natural crystals such as quartz and calcite. Other birefringent materials, for example the aforementioned liquid crystals, have nonlinear $\delta - \lambda^{-1}$ relations. One way to arrange the summation in Equation 3.18 is to sample a large number (> 100) of equally spaced points between 1000 to 1800 nm. This arrangement implies that every sampled wavelength is equally important. 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.

The merit function defined in Equation 3.18 has $n \times 2$ free parameters and is highly non-linear. A cursory investigation of this function shows that it possesses a large number of quenches, i.e., local minima. Many of the standard minimization methods for global optimization problems will fail to locate and rank these local minima. In this study, the nonlinear optimization problem of minimizing Equation 3.18 is solved by using the simulated annealing (SA) algorithm invented by Kirkpatrick et al. (1983). Title (1975) showed analytically that, using same uniaxial material, at least three layers are required in order to form a multilayer waveplate. This conclusion was reconfirmed by observing that a two-layer waveplate system will not reach a sufficiently low energy state, even if the effective temperature has been reduced to zero. For a multilayer waveplate with more than two layers, the energy reduces asymptotically as the temperature decreases.

3.3.3 Markov Chain

The driving process behind the simulated annealing algorithm is the Markov process, or Markov chain in the discrete case. A Markov chain can be uniquely described by it transition matrix **P**. Transition matrices are so-called stochastic matrices (Otten & van Ginneken 1989).

Definition 1: Given a square matrix $\mathbf{A}_{n\times n}$ whose matrix elements are $a_{ij}$where,$i, j = 1,2,\cdots,n$, $\mathbf{A}_{n\times n}$ is a stochastic matrix iff each row of the matrix sum up to 1, i.e.,

$$\sum_{j=1}^{n} a_{ij} = 1 \tag{3.20}$$

And each column summation does not vanish:

$$\sum_{i=1}^{n} a_{ij} \neq 0 \tag{3.21}$$

The matrix is called "reflexive" if $a_{ii} \neq 0$, and the corresponding Markov chain is called a "reflexive chain". ■

Definition 2: The spectrum of a matrix, $\sigma(\mathbf{A})$, is the set of all its eigenvalues, represented by: $\{\lambda_i | \lambda_i \in \sigma(\mathbf{A}), i = 1,\cdots,m\}$. ■

Stochastic matrix always has an eigenvalue equal to 1, i.e., $1 \in \sigma(\mathbf{A})$. Moreover, $|\lambda| \leq 1$. Stochastic matrix is always similar to a triangle matrix (upper or lower). The block corresponding to $\lambda = 1$ in the pseudo-diagonal form of a stochastic matrix is diagonal.

Given a transition matrix **P** of a homogeneous reflexive Markov chain, $\mathbf{P}^{\infty} \equiv \lim_{k\to\infty} \mathbf{P}^k = \mathbf{jd}^{\mathbf{T}}$, where **j** is an all 1 column vector, and **d** is the left eigenvector of the transition matrix corresponding to $\lambda = 1$. A Markov chain is reversible iff it satisfies the "detailed balance":

$$w(\mu,t)P(\mu,\nu,t) = w(\nu,t)P(\nu,\mu,t) \tag{3.22}$$

where, $P(\mu,\nu,t)$ is the transition probability from state μ to state ν at a the present time t.

Reversible Markov chain always has an ergodic set, i.e., as long as it evolves into the set, the chain will be trapped within the set while evolving.

A Markov chain which is both reflexive and reversible converges to a certain state within an ergodic set, and this state is described by the eigenvector (corresponding to the eigenvalue $\lambda = 1$) of the transition matrix $\mathbf{P}$.

These statements and definitions explain the underlying processes from the viewpoint of mathematics. However, they do not offer a direct way to solve the problem, i.e., how to construct the chain. In reality, the states such as μ or ν are so enormous that it is impossible to list all states, thus to build the transition matrix $\mathbf{P}$ and manipulate it is not possible.

In practice, a Markov chain is generated by using randomly sampling techniques, which can sample the state space more efficiently, thus reduce the number of states that need to be walked through. The technique is called variance reduction. One of the variance reduction techniques is importance sampling, in which the sampling probability of the state space is biased (non-uniformly) and later the bias is compensated when the statistical quantities are calculated.

3.3.4 Simulated Annealing

Simulated annealing is a nonhomogeneous random process composed of piecewise homogeneous Markov chains. The two adjacent chains have the following connections: (a) the uniformly decreasing controlling parameter — "temperature"; (b) the ending state of the leading chain is the starting state of the following chain. It has to be remembered that the initial state of a Markov chain does not have too much impact on the chain itself. So the connection (b) is weak.

Since the transition probability between two adjacent states of the chain is determined by the "temperature" , every single Markov chain is homogeneous, i.e., the transition matrix is not varying. The reduce of the "temperature" not only introduces a disturbance

to the current Markov chain, also, a new Markov chain starts after the reduction. Question is: how to design the reduction process so that the overall simulated annealing algorithm still converges to the global minimum? What is the constrain that need to be put on the sampling procedure and temperature reduction?

Geman & Geman (1984) proved that the scheduling has to satisfy certain condition to guarantee the global minimum to be reached. However, many practitioners do not follow the constraints provided by Geman & Geman (1984), which is stringent in the reducing rate of the temperature. In Table 3.3.4, a set of commonly used schemes of algorithm are listed. The key difference among these algorithms are the controlling of the temperature-reduction, and sampling probability. The so-called quenching algorithms are usually faster, but not guarantee to achieve the global optimized state. Practically, this can be overcome by running the program multiply times to ensure the global optimization is obtained.

3.3.5 Digression

A digression is taken here to briefly go deeper into the mechanism of the simulated annealing.

Although the Section 3.3.3 depicts the mathematical principles of the random processes involved in the simulated annealing, the discussion there did not offer any direct

Table 3.1 Examples of Simulated Annealing Algorithm

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

applicable solutions to the design of the annealing process. More advanced math tools are needed in order to be able design or guide the annealing process. One question frequently asked in the SA algorithm is when the equilibrium has been reached with each Markov chain? Classically this is done by measuring certain typical quantities of the system, (such as the internal energy, or the average) at adjacent time intervals and compare. This primitive routine can be avoided if the converging rate of the Markov chain can be determined in advance, so that a threshold can be set to determine the number of the steps needed to reach this threshold. The determination of converging rates of a random process is one of the major subjects of the Large Deviation Principle (LDP). The LDP was applied to the random walk by Mogulskii (1974) and to the Brownian motion by Schilder (1966). However this subject needs extensive efforts to explore.

3.3.6 Different Thicknesses, Different Azimuths

The most general case of the multilayer achromatic waveplate is that the thickness and azimuth of every layer are controllable variables. The optimization results are show in Table 3.2. Note that in the three-layer $\lambda/4$ waveplate configuration in Table 3.2, 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 an 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 3.18, the error of the computed configurations are "globally" optimized from 1000 to 1800 nm. The analytical configurations were optimized locally. In the following section, the optical properties of the configurations listed in Table 3.2 are discussed based on several computational setups.

Table 3.2 Configurations of the n-Layer Achromatic Waveplates

Type	#	θ_i	δ_i	θ_i	δ_i	θ_i	δ_i
	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″
$\lambda/4$ waveplates	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″
	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″
$\lambda/2$ waveplates	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, $\delta \propto \lambda^{-1}$ is assumed, which is true for natural crystals such as quartz and calcite. Note that throughout the paper, only the properties of $\lambda/4$ waveplate are discussed in details. $\lambda/2$ waveplate can be analyzed in a similar way.

3.4 Properties of Multilayer Achromatic Waveplates

Numerical methods can be helpful before lab experiments can be done. Several important concerns about the waveplates are: the achromatism; the sensitivity to the light beam speed; the stability of the equivalent fast axis.

3.4.1 Normal Incidence

Assume the incoming monochromatic light incidents perpendicularly. In the first experiment, waveplates are set up between two linear polarizers, see the experimental arrangement in Figure 3.3 (a six-layer achromatic $\lambda/4$-waveplate is shown in the figure.). 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. If the waveplates are assumed to be perfect $\lambda/4$ waveplates, the output intensity shall be identical as the second polarizer is rotating. However, 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 through multilayer $\lambda/4$ waveplate is elliptically polarized (see Figure 3.4 which is plotted for polarized light passing through a single crystal, three-layer, six-layer, and ten-layer $\lambda/4$ waveplates). 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. At these locations, the multilayer waveplate resembles a perfect $\lambda/4$ waveplate. 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 previous studies (Ma et al. 2004; Beckers 1971; Beckers et al. 1975), since the effective optical axis of a three-layer achromatic waveplate rotates to different angles at different wavelengths. 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.

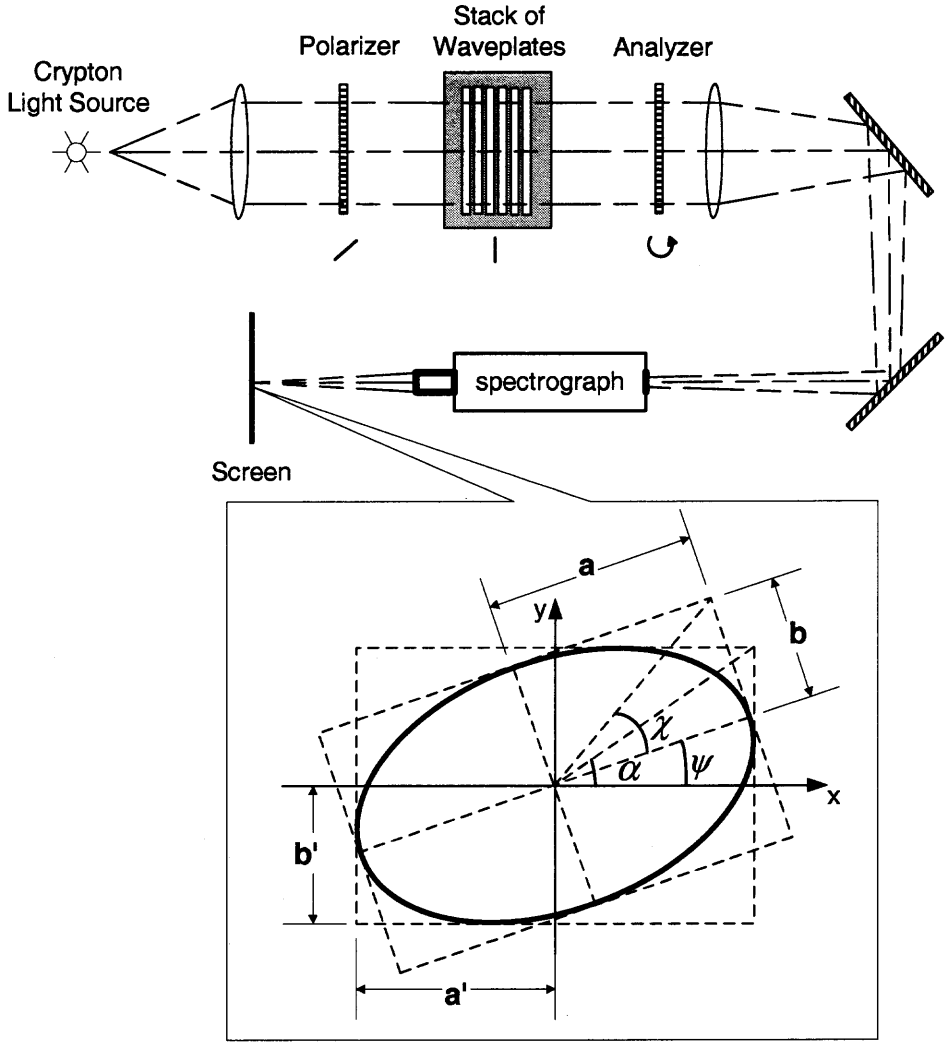

Figure 3.3 Measurement of retardance.

Given the ellipticity of the testing light obtained in the above mentioned method, the retardance can be derived by using the method shown in Figure 3.3. The corresponding results are plotted in Figure 3.5, 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 0.3% wave, or equivalently, 1°. As will be discussed in later sections, the retardance error in this magnitude is negligible in the application of a birefringent filer or a phase shifter.

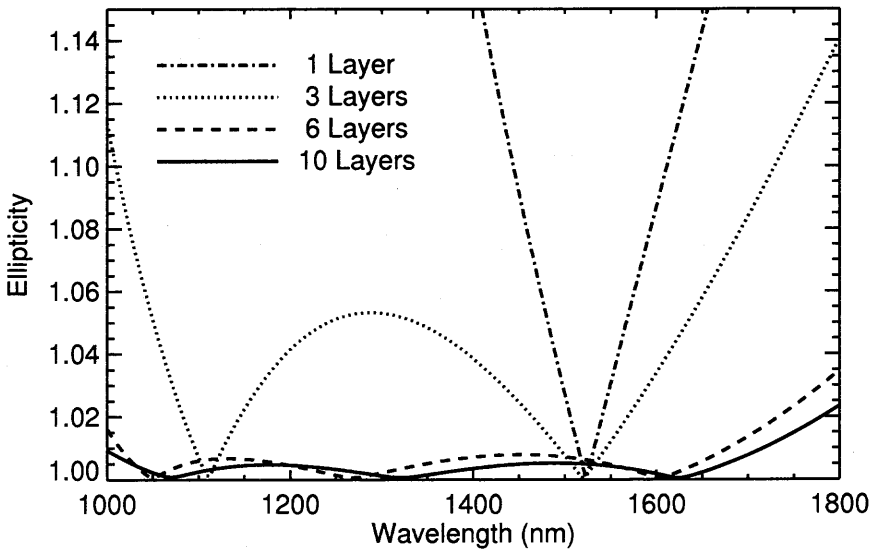

Figure 3.4 Ellipticity vs. wavelength.

3.4.2 Polarization Transformation in n-layer Waveplates

The second experiment is concerned with the polarization transformation properties of n-layer achromatic waveplates. Let the incoming light be linearly polarized. It can be shown that the Stokes vector of linear polarized light along a direction α with respect to the x-axis is

$$\mathbf{S} = [1, \cos^2\alpha - \sin^2\alpha, 2\sin\alpha\cos\alpha, 0]^T \tag{3.23}$$

Let $\mathbf{S}_{\text{in}}$ 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 $\mathbf{M}_{\text{awp}}$ is the Mueller matrix of an n-layer achromatic waveplate. In Figure 3.6 (see Table 3.2 for parameters of each layer). The two vertical straight lines indicate the locations of wavelengthes 1083 nm and 1565 nm. The effective optic 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 $\pi/5$. The corresponding Stokes vector of the light is $[1, \cos^2\frac{\pi}{5} - \sin^2\frac{\pi}{5}, 2\cos\frac{\pi}{5}\sin\frac{\pi}{5}, 0]$. This computational simulation shows that, after passing through a $\lambda/4$ waveplate, polarization energy is re-distributed among Q, U, V-components and will cause varying polarization

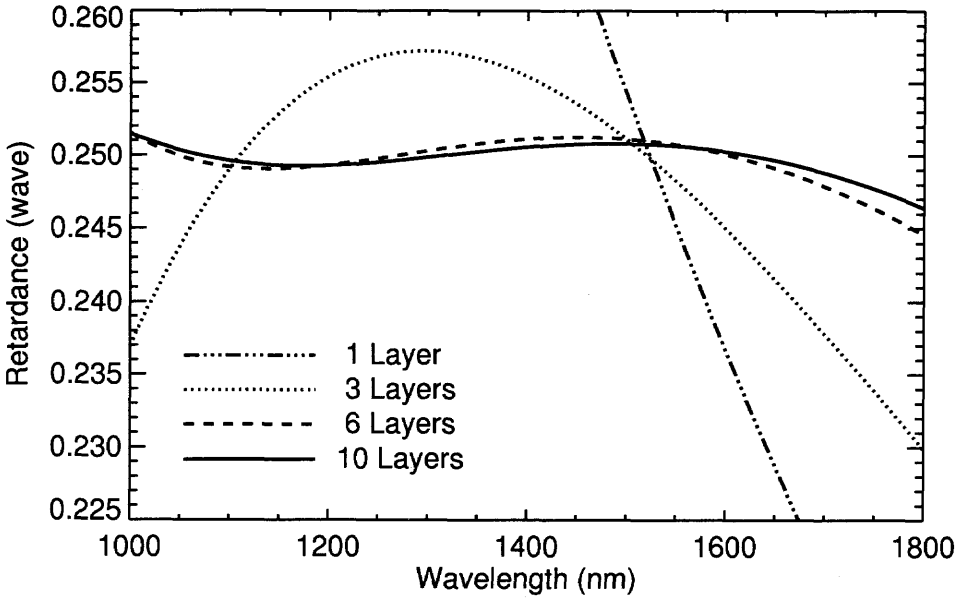

Figure 3.5 Retardance vs. wavelength (computational), derived from Figure 3.4.

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 transversal magnetic field. The V-component is related to the longitudinal field. Therefore, the crosstalk of V-component due to the polarization error of waveplates can appear as spurious magnetic signal in Q-component. This was observed before — in the calibration of DVMG system at BBSO, 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).

In Figure 3.6, the crosstalk due to n-layer waveplate is roughly limited to 2% across the specified wavelength range. 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 Muller matrices need to be investigated for distinctive wavelengths and for different azimuth. Only the elements $\mathbf{M}(\mathrm{i,j})$ with $i, j = 2,3,4$ are considered, which define the crosstalk between linear polarizations and circular polar-

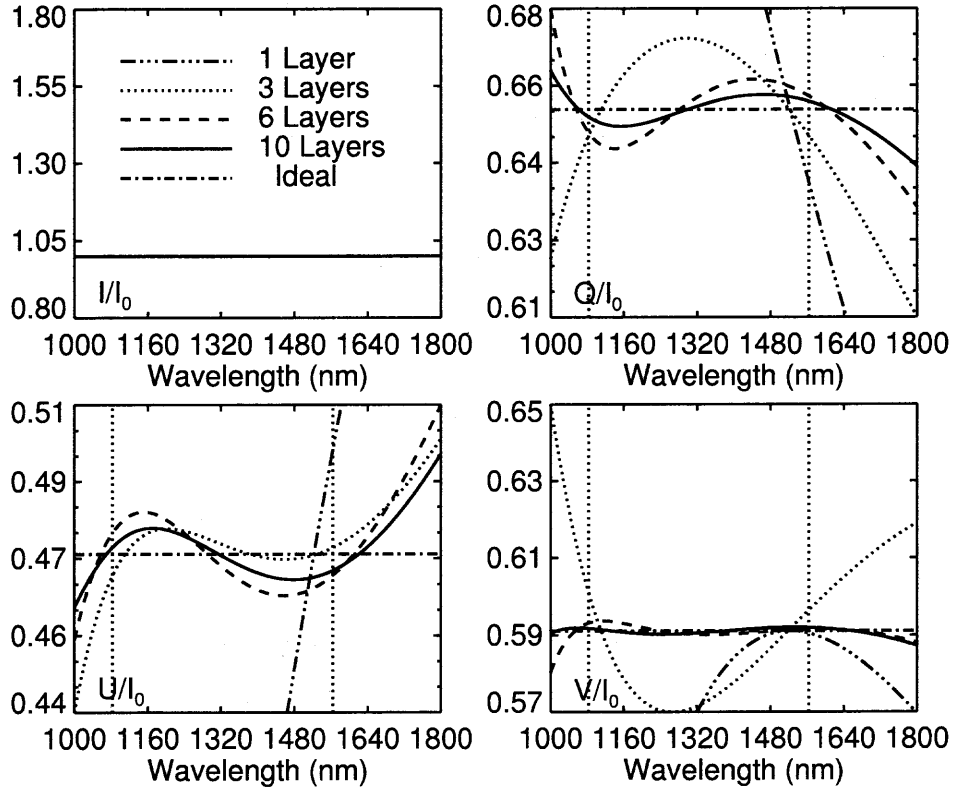

Figure 3.6 Simulated polarization crosstalk of n-layer $\lambda/4$ waveplates.

izations. Let

$$\Delta\mathbf{M} = \mathbf{M}_{10} - \mathbf{M}_{\text{ideal}} \tag{3.24}$$

be the difference between the Mueller matrix of a ten-layer quarter waveplate ($\mathbf{M}_{10}$) and an ideal quarter waveplate ($\mathbf{M}_{\text{ideal}}$). In Figure 3.7, the errors of the Mueller matrices elements of a ten-layer achromatic $\lambda/4$ waveplate with respect to an ideal $\lambda/4$ waveplate at a single wavelength $\lambda_0 = 1083$ nm. Both waveplates are rotating from 0°~360°. Scale of all y-axes is 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.

Two features can be read from this figure: (1) the magnitudes of all the elements of $\Delta\mathbf{M}$ are oscillating with respect to the azimuth of the rotating waveplate, except $\Delta\mathbf{M}(4,4)$; and (2) 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.

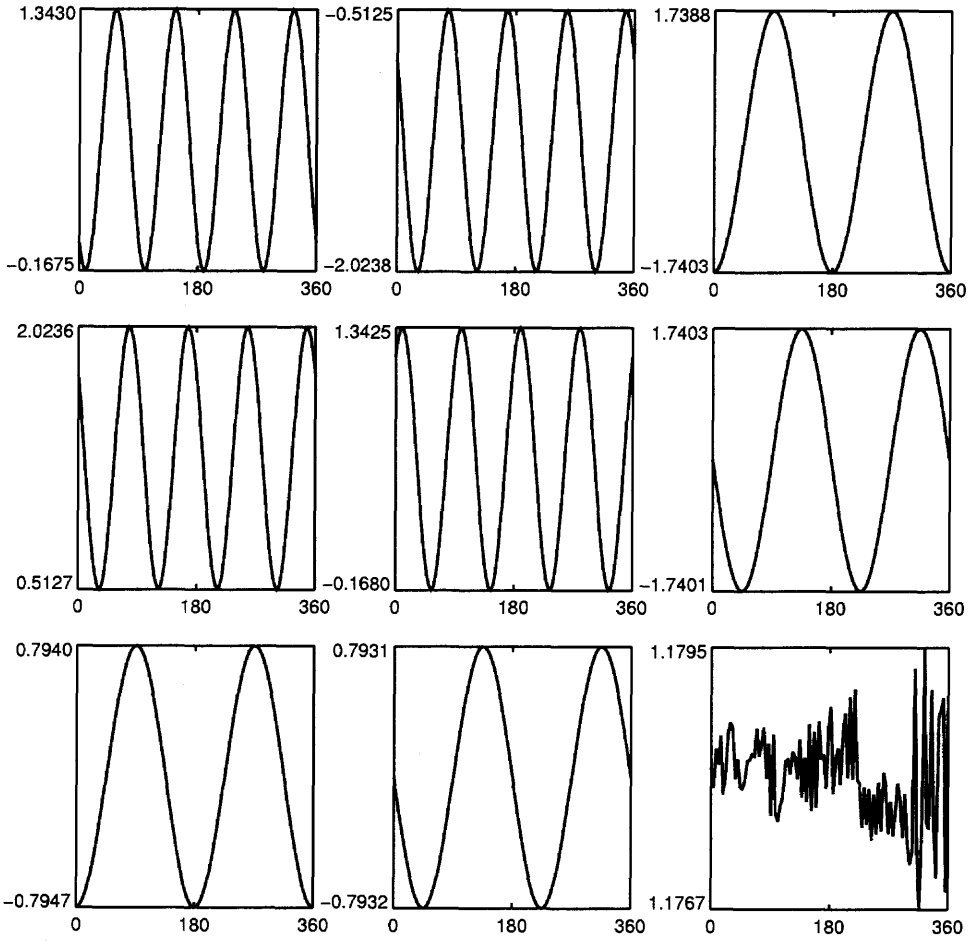

Figure 3.7 The errors of Mueller matrices elements.

3.4.3 Orientation of the Optical Axes of n-Layer Waveplates

The optical axes of Pancharatnam-type achromatic waveplates (Pancharatnam 1947) rotate to distinctive angles at different wavelengths (Beckers et al. 1975). This issue is minimized in six- and ten-layer achromatic waveplates. In the criteria of the optimization problem defined by Equation 3.18, $\Delta\mathbf{J}$ 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, which is zero. In Figure 3.4 and Figure 3.7, it is also shown indirectly that the azimuth of optical axes of six- and ten-layer achromatic waveplates do not change severely within the specified wavelength range. Direct evaluation of the azimuth of optical axis involves the decomposition of complex Jones matrices and other complicated algebra operation. An easier way to evaluate the optical axis is doing lab experiments on these multilayer waveplates.

3.4.4 Off-Axis Effects

Let α be the incident angle, which describes the off-axis effect. For thin crystal slab, the refraction effects on the interfaces are neglected. Also the convention is adopted 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 can be written as (Evans 1949)

$$\delta = \delta_0 \left[1 + \frac{\alpha^2}{2n_o} \left(\frac{\cos^2 \beta}{n_e} - \frac{\sin^2 \beta}{n_o} \right) \right] \quad , \tag{3.25}$$

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 greater 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 3.25 has to be replaced by $\beta + \pi/2$. Consider using quartz as the material of making the n-layer waveplates. The refractive indices of crystal quartz at different wavelength can be 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 3.2 is approximately $1'\ 30''$ according to Equation 3.25.

3.4.5 Application - Phase Shifter

Phase shifting unit is an instrument that produces a continuous fractional phase difference between two perpendicularly polarized light. Evans (1949) discussed several approaches theoretically to realize a phase shifter. One of these approaches consists of a combination of one $\lambda/2$ waveplate sandwiched by two $\lambda/4$ waveplates, which is particularly interesting due to its flexibility. This approach is discussed by substituting the normal birefringent crystals with 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 both the first waveplate unit $(\lambda/4)$ and the third waveplate unit $(\lambda/4)$ are set to 45° with respect to the polarization direction of the incident light, without loss of generality, it is chosen to be the x-axis. The second waveplate $(\lambda/2)$ is sitting on a rotary stage driven by a motor and its azimuthal angle θ is measured with respect to the x-axis. In Figure 3.8, elements of the Jones matrices of phase shifters are examined. The retardance of a phase shifter is derived from the two diagonal elements of the Jones matrix, see Figure 3.8a. In Figure 3.8b, the amplitude of one anti-diagonal element (complex number) is plotted. The solid curve represents the phase shifter made of six-layer achromatic waveplates, which are almost zero. As a comparison, the Jones matrix anti-diagonal elements of a phase shifter made of three single-crystal waveplates (dashed curves) are overplotted on the same figure. 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 3.8a also shows that the retardance of the phase shifter made of achromatic waveplates is linearly related to the azimuth 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. 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 a phase shifting rate of 8 waves/s, i.e., 125 ms/wave. Phase shifters can also be made from liquid crystal, i.e., Liquid Crystal Variable Retarders (LCVR). Meadowlark LCVR response time to a half wave voltage is 10 ms according to the commercial documentations. 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. The advantage of using an 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).

3.5 Laboratory Experiments

3.5.1 Measurement of Birefringence

The scheme shown in Figure 3.3 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. 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^2 and b^2 respectively in Figure 3.3. The ratio of a/b is the ellipticity shown in Figure 3.4. Notice that another commonly used measure of the elliptical shape, eccentricity ε, is related to the ellipticity by $\varepsilon = \sqrt{1-b^2/a^2}$.

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 3.3 can be easily obtained (see, for example, Born & Wolf 1999):

$$\sin 2\chi = \sin 2\alpha \sin \delta \tag{3.26}$$

$$\tan 2\psi = \tan 2\alpha \cos \delta \tag{3.27}$$

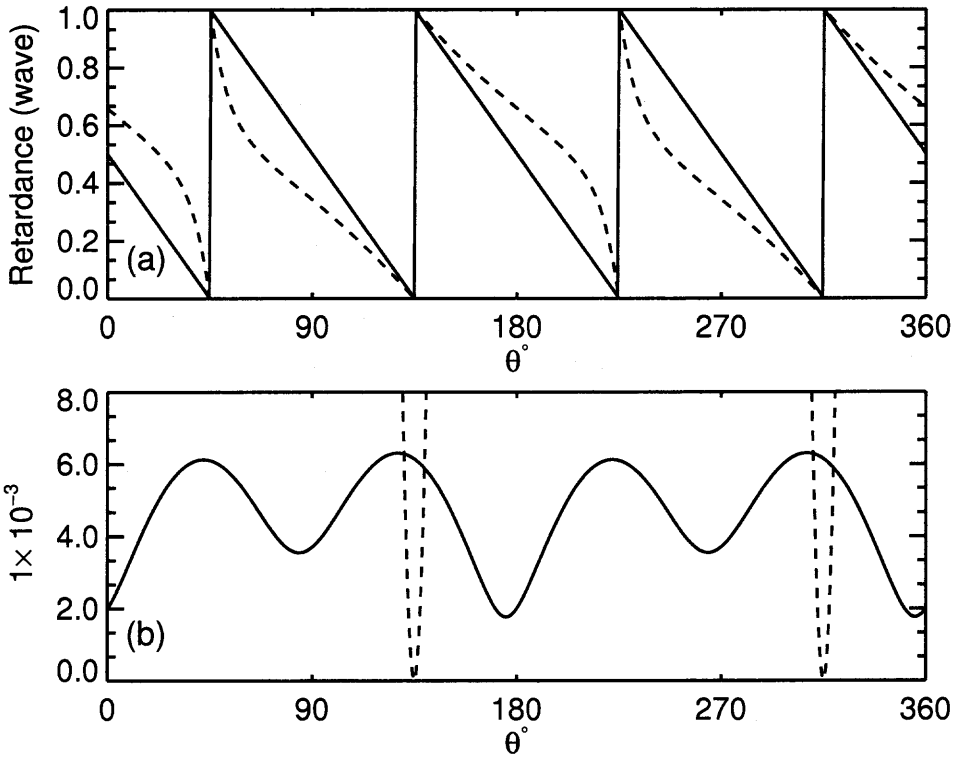

Figure 3.8 Evens three-waveplate phase shifter (Evans 1949) using achromatic waveplates (solid curves). (a) Retardance vs. θ (azimuth of the half waveplate in the middle) and (b) the magnitudes of the anti-diagonal elements of the Jones matrices of the two phase shifters. The dashed curves refers to 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. The wavelength of the incident light is $\lambda = 1083.3$ nm.

from which, it can be shown

$$\sin 2\alpha = \sqrt{\left(\sin^2 2\chi + \tan^2 2\psi\right) / \left(1 + \tan^2 2\psi\right)} \tag{3.28}$$

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,

$$\tan\chi = \mp b/a \tag{3.29}$$

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:

$$0 < \psi < \pi/2 \quad \text{if} \quad 0 < \delta < \pi/2 \tag{3.30}$$

$$\pi/2 < \psi < \pi \quad \text{if} \quad \pi/2 < \delta < \pi \tag{3.31}$$

In the case of using Equation 3.26 to derive δ, it can be shown:

$$\delta = \sin^{-1}\left(\frac{\sin 2\chi}{\sin 2\alpha}\right), \text{if} \quad 0 < \psi < \pi/2 \tag{3.32}$$

$$\delta = \pi - \sin^{-1}\left(\frac{\sin 2\chi}{\sin 2\alpha}\right), \text{if} \quad \pi/2 < \psi < \pi \tag{3.33}$$

These two relations were implemented in deriving the plots in Figure 3.5.

3.5.2 Waveplates For IRIM

Another example that needs broadband achromatic waveplates is the Near Infrared Imaging Magnetograph (IRIM) at Big Bear Solar Observatory (Denker et al. 2003a,b). The IRIM system utilizes a Fabry-Pérot interferometer, two birefringent filters, and a set of wide band interference filters. The two birefringent filters are designed for He I 1083 nm and Fe I 1565 nm respectively due to the chromatism of optical components, e.g., waveplates, polarizers and coating. Optical components of the 1083 nm birefringent filter were designed to be removable, and therefore normal waveplates can be easily substituted by achromatic waveplates. The achromatic design of IRIM is challenging due to the two lines are far away apart. In order to solve this problem for IRIM, the merit function is redefined as

$$E(\phi_k, \theta_k) = \sum_{\Delta\lambda_1}\sum_{ij} w_{ij}\Delta J_{ij}\cdot\Delta J_{ij}^{*} + \sum_{\Delta\lambda_2}\sum_{ij} w_{ij}\Delta J_{ij}\cdot\Delta J_{ij}^{*} \quad , \tag{3.34}$$

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. $\mathbf{J} = \mathbf{J_k}$ is the Jones matrix of the k-layer waveplate. As shown 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. The simulated annealing program is ran for this newly defined merit function, and the following solution (in radians) are found:

$$\theta_1 = 0.53429 \quad \delta_1 = 1.69690,$$
$$\theta_2 = 2.45617 \quad \delta_2 = 2.63967,$$
$$\theta_3 = 0.53423 \quad \delta_3 = 1.69677.$$

A numerical simulation is carried out for the third stage of the Lyot filter used in IRIM with the same design parameters of the filter calculated by Wang et al. (2001). In this simulation, the normal single-layer waveplates (half and quarter) are substituted by both three-layer waveplates designed above and six-layer waveplates in Table 3.2. The transmission profiles were calculated specifically in the vicinities of 1083 nm and 1565 nm. And the rotating half-waveplate is adopted as the bandpass tuning mechanism. Two snapshots are taken when the azimuth of the rotating waveplate are 0 and $\frac{2}{7}\pi$. The corresponding transmission profiles of a single stage Lyot filter are presented in Figure 3.9.

There are two improvements to the bandpass profile of a birefringent filter by using n-layer achromatic waveplate — the bandpass symmetry and the off-band scattered light level (Figure 3.9). To summarize,

(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. In the right panel of the second row in Figure 3.9, 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. The periodicity of the symmetry of the bandpass are investigated, and it was 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 matrix errors of three-layer waveplates were minimized at 1083 nm and 1564 nm by simulated annealing procedure. It is expected to see the bandpass shape in the vicinities of these two wavelengthes should be close to ideal achromatic waveplates. On the contrary, it can be seen that the scattered light level is still very high. One known source of error is the wavelength-dependency of the fast-axis of three-layer waveplates (Beckers et al. 1975).

(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, the following conclusions can be drawn regarding the application of *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. Exercises using lower number layers than six are not tried except for three, 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. Using the solutions of multilayer waveplates provided in the current paper, the filter design can be greatly simplified.

3.5.3 Experiment Results Discussion

A set of multi-layer waveplates were manufactured by Nanjing Institute of Astronomical Instrumentation (China), according to the parameters shown in Table 3.3. This set of values for achromatic waveplates belong to earlier version of the design, and the details of the design such as the reference wavelength, is missing. However, these are the experimental data available for the current writing of the thesis.

Two calibrated near infrared linear polarizers are used in two combinations: crossed and parallel. The sample of waveplates are inserted in between with an angle 45° to the active axis of the first polarizer. Note that a half waveplate can rotate linear polarized light by 90° if the relative angle between the light and the optic axis of the waveplate is 45°. Given the waveplate inserted is a perfect half waveplate, the output intensity should be maximum if the polarizers are crossed, and be minimum if the polarizers are parallel. The three-layer quarter waveplate was not tested as a quarter waveplate. Instead, two such quarter waveplates were combined together and tested as a half waveplate. Therefore, this test can only prove the three-layer waveplate is an half waveplate or quarter waveplate, but did not measure the retardance.

The testing results of a set of three-layer waveplates are plotted in Figure 3.10, 3.11, 3.12, and 3.13. It can be seen from Figure 3.12 and Figure 3.13 that the wavelength dependence of the waveplate reduced dramatically, if not perfectly. This test proves that the three-layer waveplates do approach the achromatic waveplate. Increasing the number of layers is a promising direction to solve the achromatism of waveplates.

3.6 Conclusions

In this study, a numerical procedure is developed based on a Monte Carlo simulate annealing algorithm to design crystalline multilayer achromatic waveplates. The procedure was applied to three examples of multilayer waveplate, three-, six- and ten-layer waveplates, and obtained the parameters given by Table 3.2. Either six- or ten-layer structures

can produce high quality $\lambda/4$ and $\lambda/2$ achromatic waveplates in wide wavelength regions. Specifically, the optical-axis issue in three-layer waveplate is minimized in six- and ten-layer waveplates (see Figure 3.4). The performance of such n-layer achromatic waveplates was numerically evaluated by considering the residual errors in Jones/Mueller matrix elements and polarization errors in Stokes vectors. The impact of off-axis light is briefly considered. Using the six- or ten-layer achromatic waveplates, an achromatic phase shifter can be constructed based on a simple structure proposed by Evans (1949). This phase shifter continuously produces fractal phase change from 0 to 2π within the spectral range considered. Therefore, a reliable tuning solution can be found for many instruments, particularly, in our case, for the tunable birefringent filters.

It might turn out that thin material with stress birefringence (photo-elastic effect (Born & Wolf 1999)) is one of the possible choices other than quartz or calcite. Wafers of stress 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, care should be taken when using these non-crystalline materials. The reason is that the parameters shown in Table 3.2 are designed based on the assumption that the "retardance vs. wavelength" relation is Equation 3.19. This relation need to be re-evaluated for non-crystalline materials. In these cases, the exact $\delta - \lambda$ function has to be known. However, the SA algorithm is sufficiently flexible to allow the implementation of different $\delta - \lambda$ functions.

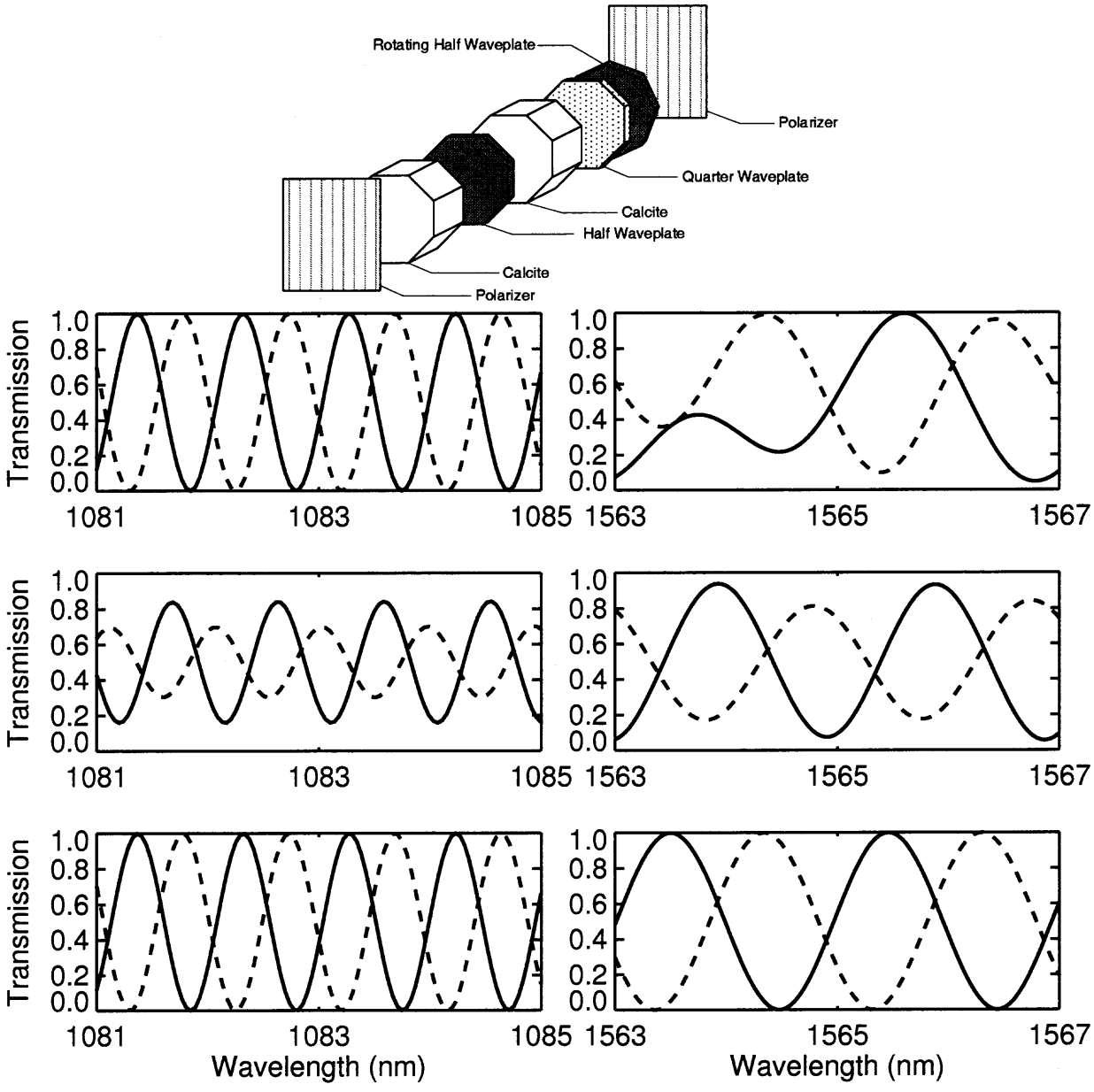

Figure 3.9 Tuning of a single stage of a Lyot filter - single stage. Solid curves: azimuth of the rotating $\lambda/2$ waveplate is set at 0; dotted curves: azimuth of the rotating $\lambda/2$ waveplate is set to $\frac{2}{7}\pi$. First row: a cartoon of a single stage; second row: transmission of the Lyot stage using single crystal waveplates; third row: using three-layer achromatic waveplates; forth row: using six-layer achromatic waveplates.

Table 3.3 Achromatic Three-Layer Waveplates

Half AWP		Azimuth	Thickness (mm)
	1 (thick)	60° 9′59″	0.5755 (0.595)
	1 (thin)	150° 9′59″	0.5000 (0.519)
	2 (thick)	118°20′53″	0.5755 (0.595)
	2 (thin)	28°20′53″	0.5000 (0.519)
	3 (thick)	60° 9′59″	0.5755 (0.595)
	3 (thin)	150° 9′59″	0.5000 (0.519)
Quarter AWP			
	1 (thick)	62°55′ 4″	0.5617 (0.561)
	1 (thin)	152°55′ 4″	0.5000 (0.499)
	2 (thick)	135°25′22″	0.5924 (0.610)
	2 (thin)	45°25′22″	0.5000 (0.519)
	3 (thick)	62°55′ 4″	0.5617 (0.561)
	3 (thin)	152°55′ 4″	0.5000 (0.499)

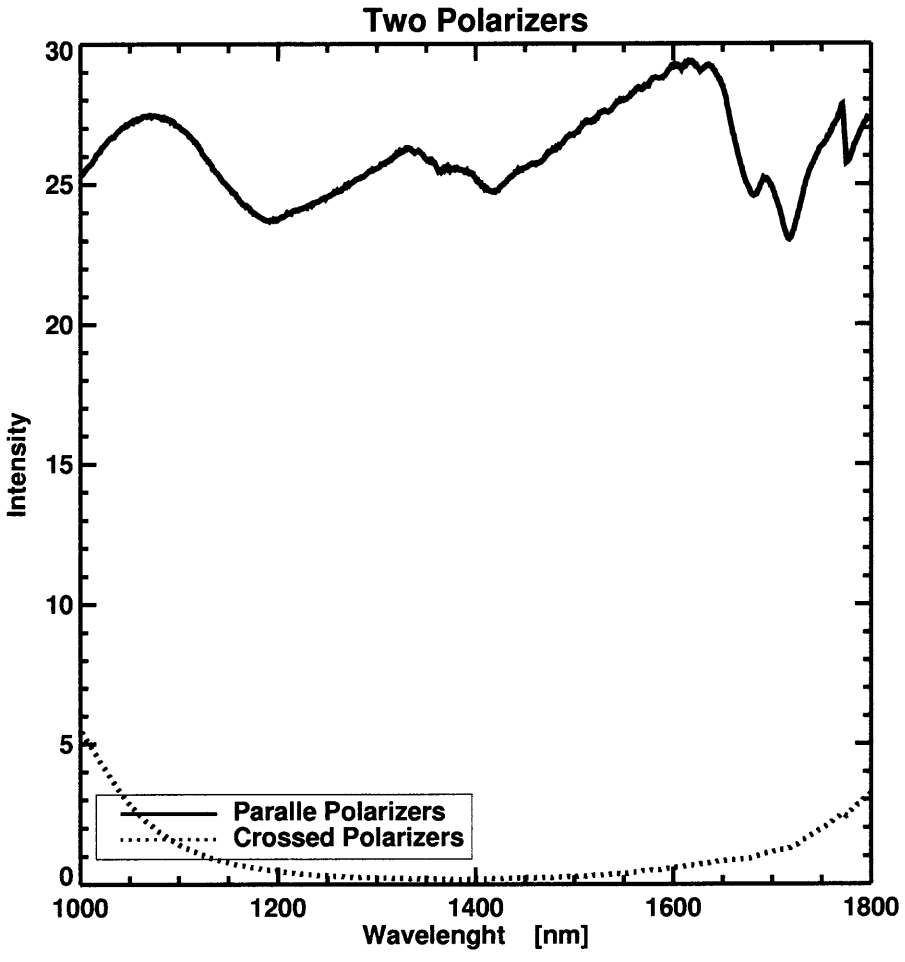

Figure 3.10 Properties of the near infrared linear polarizers: crossed and parallel. Also, the profile of the light source is involved in these profiles.

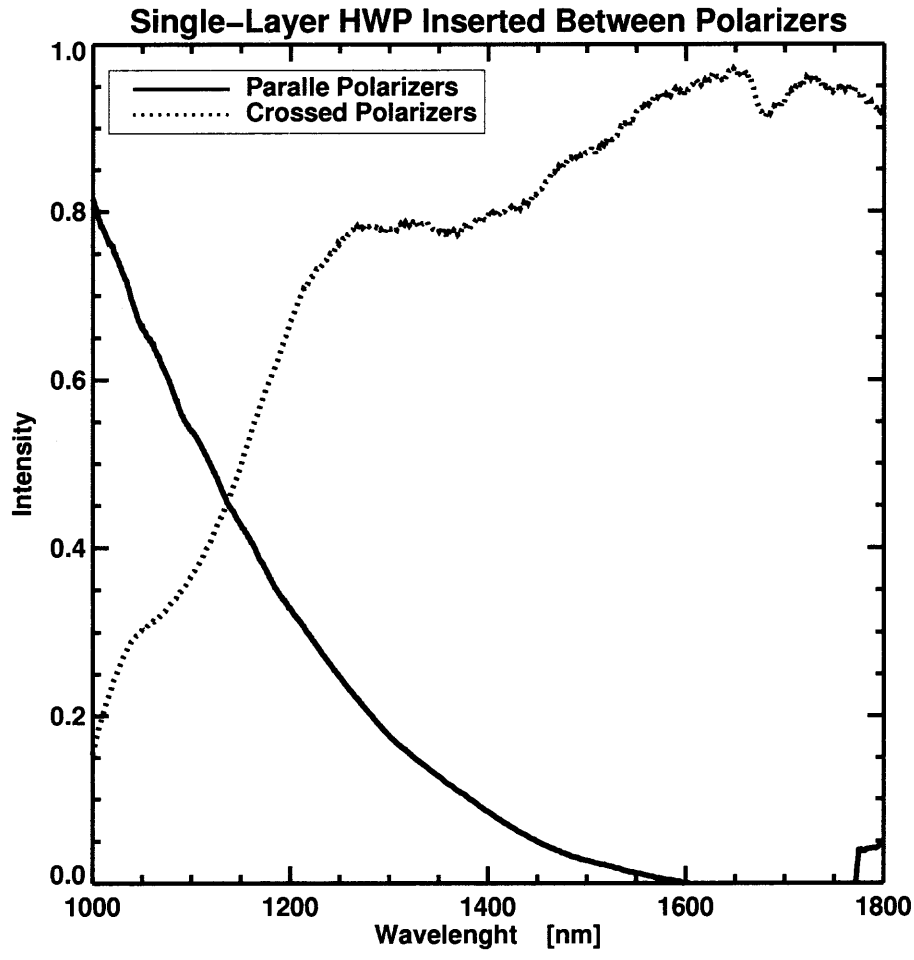

Figure 3.11 Testing results for a normal single-layer half waveplate using the polarizers and light source mentioned in Figure 3.10.

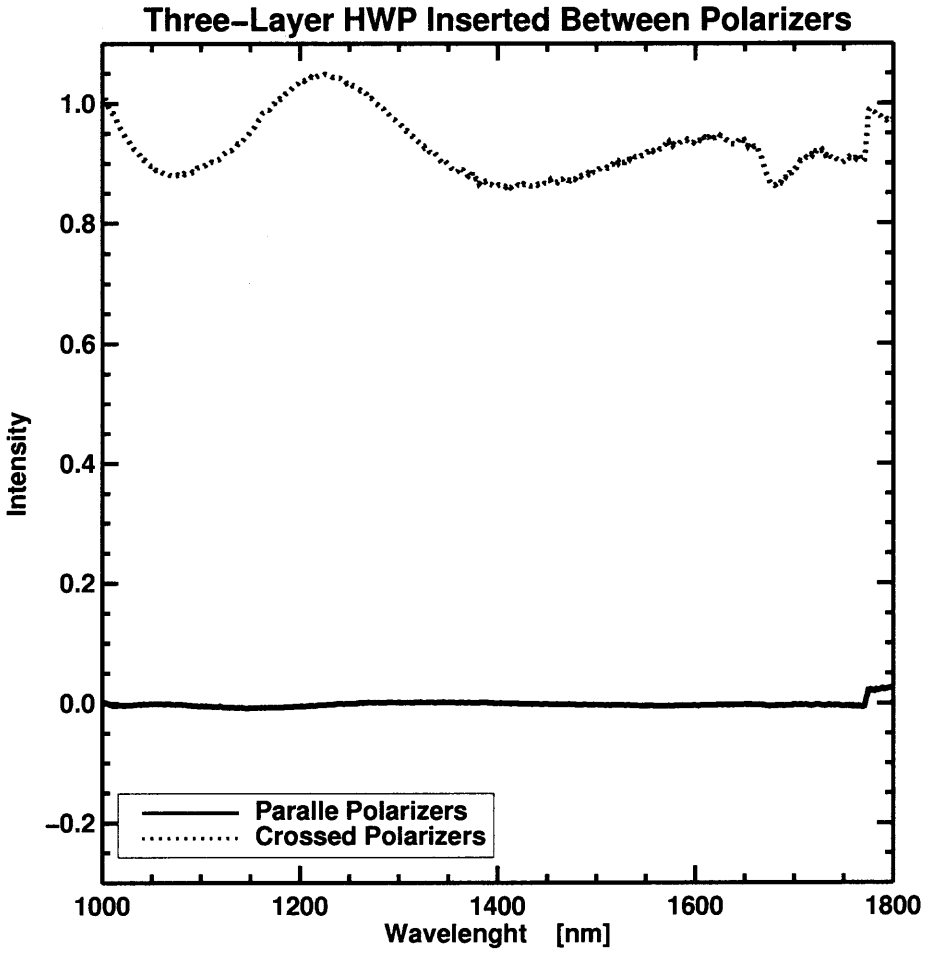

Figure 3.12 Testing of a three-layer half waveplate in Table 3.3.

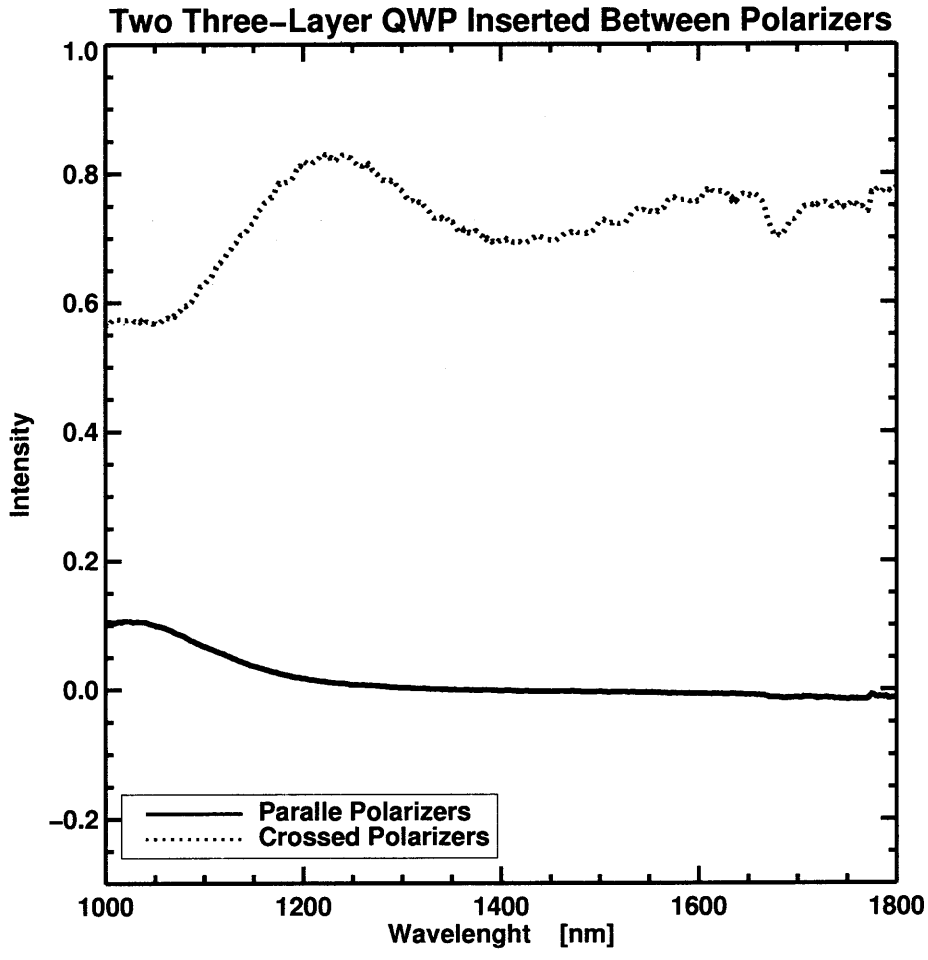

Figure 3.13 Testing of the two adhered three-layer quarter waveplates in Table 3.3.

CHAPTER 4

A LARGE SUNSPOT AT 1.56 μm CONTINUUM

4.1 Motivations

Infrared observations of sunspots have the following advantages: 1) Zeeman splitting provides three times the sensitivity at 1.56 μm as at 0.5 μm; 2) Stray light is less of a problem in the infrared for the reasons that instrumental stray light is usually low, and the effects of stray light within sunspots are small due to the intrinsic intensity contrast is lower; 3) Opacity of the solar atmosphere reaches minimum at 1.6 μm, therefore observations at near infrared discover the sunspot structure on the lowest visible layer (Bruls et al. 1991); 4) It is possible to measure a nearly true continuum intensity and make a straightforward estimate of temperature (Maltby et al. 1986). More specifically, there is a close correlation between the continuum brightness at 1.56 μm and the kinetic temperature at the height of line formation (Vernazza et al. 1976). Therefore, the temperature can be derived from the continuum brightness using black body radiation — Planck's function.

4.2 Observation

The set of data used in the current chapter was taken at 76 cm Dunn Solar Telescope of National Solar Observatory (DST/NSO) with the high-order adaptive optics system. It enabled us to acquire diffraction-limit data sets (0.″14 at G-band 430 nm). The time range presented here is from 10:04:59 UT to 10:31:58 UT on December 2nd 2004. A tunable near infrared birefringent filter (Wang et al. 2001, 2004) was set up in front of a Rockwell camera with 1024$\times$1024 CMOS chip. The passband was shifted to 1.56 μm continuum to avoid any solar lines. The lab test at Cambridge Research & Instrumentation Inc. (CRi) showed the band width of the Lyot filter in the vicinity of 1.56 μm is 0.22 nm with a peak transmission close to 35%. This Lyot filter along with a near infrared pre-filter allow the

exposure time to be set as around 1/10 sec in near infrared. An image selection algorithm was applied to save one best frame out the ten. The image scale of the data set is 0″.12 per pixel.

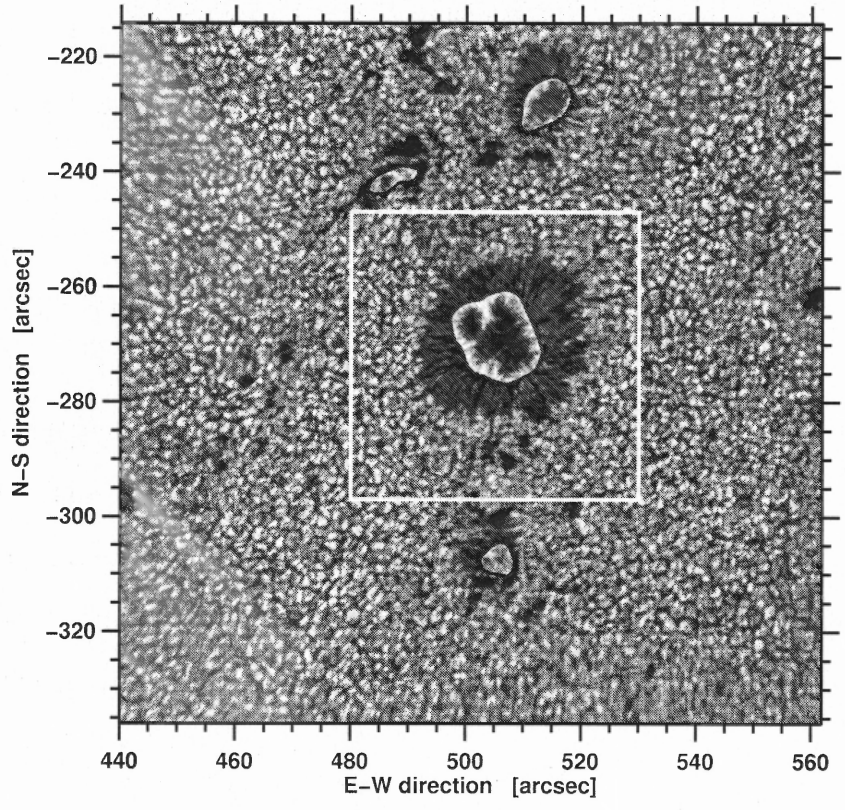

Figure 4.1 Region of interests shown on the full FOV of the observation.

Total 996 frames were recorded for the active region NOAA 10707 locate at S17°W24°, see Figure 4.1. This region is composed of one large sunspot (about 32″in diameter) , one small sunspot, and several pores. The large sunspot is the target of study in this chapter.

Figure 4.2 shows the RMS contrast of the observation with high-order adaptive optics at NSO (Rimmele et al. 2003). RMS contrast is defined as:

$$rms. = \frac{\bar{\sigma}}{\mu} \tag{4.1}$$

where $\bar{\sigma}$ is the standard deviation of the image intensity; μ is the mean intensity of the

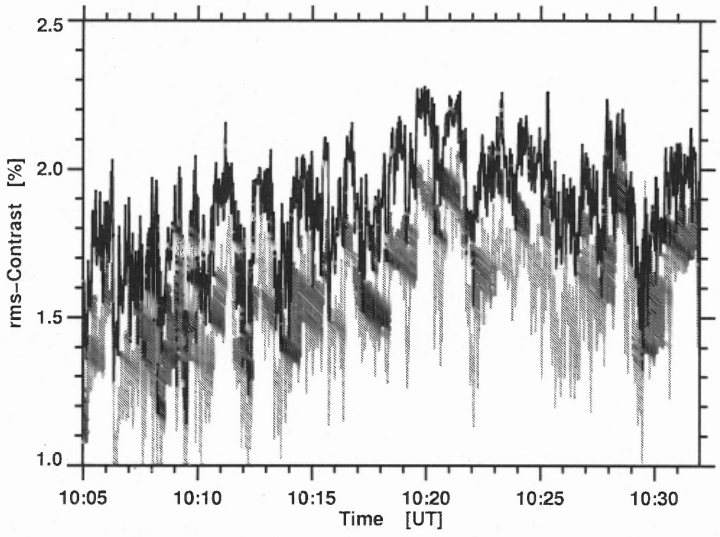

Figure 4.2 RMS contrast of the observation.

image. Standard deviation evaluates the fluctuation of intensity, which can be biased by global illuminating. Divided by the mean intensity, such global illuminating can be suppressed, which make it a more objective parameter to evaluate the quality of the image. As regards to the techniques processing the global illuminating, subsonic filter is used in the later analysis.

The darker curve shows the RMS contrast evaluated at the center granulation region; lighter curve shows the granulation at the edge of FOV. Since the high-order adaptive optics system correct the wavefront distortion according to a reference feature in FOV (in this case, the large sunspot), it is expected that the neighboring region of the sunspot is corrected more accurately than the regions further away from the sunspot.

4.3 Data Reduction and Processing

After a routine correction for dark current and flat fielding, the sequence of images were filtered using subsonic filter (Title et al. 1989) ($v_{ph} = 4$ km/s) to remove 5-minute oscillation. Only 200 frames are kept after the filtering. Subsonic filtering defines a cone (u,v,t) in spatial-temporal frequency space instead of in (u,v) only.

UDs are identified by a region growing algorithm, see Figure 4.3. The first image is enhanced by an edge sharpening routine for better demonstration; the second image is the ROI masked out from original image; local maxima are detected and highlighted in the third image; finally, UDs detected by a region-growing routine are shown in the last image. Only those regions enclosing more than four pixels are considered as a real UD. Accordingly this sets the sensitivity of the detection algorithm. Therefore, not every highlighted point in the third image has a counterpart in the last image. The following steps are applied to find UDs: (1)Identify the center of UD. Identified the brightest points of UDs' (every local peak among a 3 by 3 square is recognized as the center of a UD); (2) Grow a UD region. Starting from this brightest point, the second-derivative is calculated for each pixel recursively; (3) Determine the boundary of a UD. There are total 4 components in this derivative: NS, EW, NW, SE. As long as at least one of these 4 components is greater than zero, the pixel is recognized as the boundary of current UD.

This algorithm is always "loose" in the sense that only one positive derivative component is required. Rigorously, derivative component along radial direction centering at the brightest point should be positive for boundary pixel. However, due to the noise in the image, algorithm based on radial derivative can be unreliable. The chance of running into a pre-mature fake boundary pixel is same for both algorithms.

Although the adaptive optics improved the quality of images, the UDs observation require even higher image quality. For this reason, only those frames with rms contrast

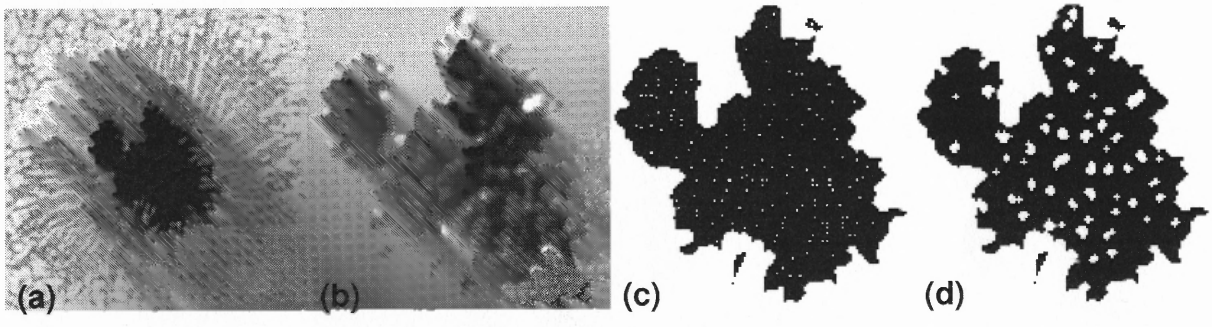

Figure 4.3 ROI (271px, 32″square) and image processing. The image shown in the figure is taken at UT 10:21:37 (frame# 607/996).

(neighboring granulation regions) greater than 2% are selected for the purpose of size and intensity measurements.

4.4 Sunspot Decaying of NOAA 10707

An intensity threshold of $0.8I_{ph}$ for umbral region (Figure 4.3) is chosen visually to include almost all the umbral region, and exclude the penumbra meanwhile. The planar area of the umbral region is 1.67×10^6 Mm^2 $\sim1.96\times10^6$ Mm^2 during the observation period. The umbral region shrunk about 6% in the end of the observation nonlinearly. Due to the limited time span of the observation, later phase of the active region was not clear.

4.5 Fine Structures Inside of Sunspot

Within the penumbral region, the filamentary structures appear to experience a splitting — an elongated penumbral grains not only move towards umbra, also outer part of the penumbral is moving outward. A direction reverse of the proper motion happens accordingly, see Figure 4.4 and Figure 4.5. The penumbral grains moving inward also have a Gaussian-like brightness distribution. However, this does not apply to outward-moving penumbral grains, many of which move similarly as smoke in the air blown away from umbral region — streaming with twisting.

The strong light bridge intruding from N-E sends plasma stream into the umbral region with a speed as fast as 1 km/s (fastest speed appears near the boundary of light bridge and umbra.). The intensities of all observed penumbral grains reduce after they move into umbra, some of them disappear at the boundary of penumbra and umbra.

Inside the umbra, two filamentary links of bright dots can be confidently identified. During our 27-minute observation, these two links experience only minor changes with regard to the size of bright dots composing the links and the relative positions between dots. The region between the two links are the darkest area which has lowest intensity of

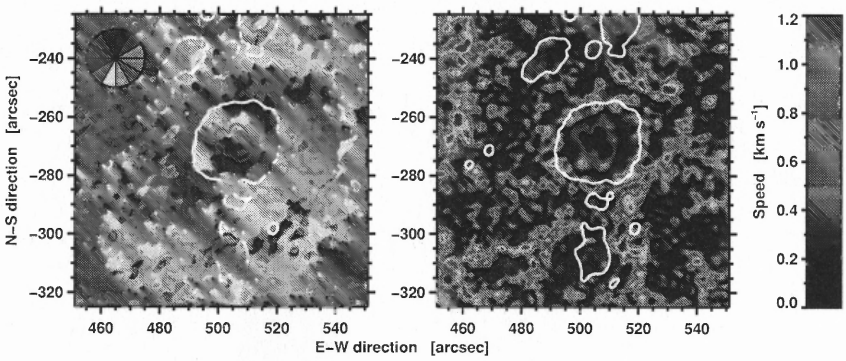

Figure 4.4 (Color) Proper motion of the FOV of the observation. Left: the flow direction; Right: the magnitude of velocity. Note that only the large sunspot at the center is of special interests with this presentation.

$\sim 0.6 I_{ph}$ (umbral nucleus). The umbral region in the E-N is almost isolated by intruding light bridge, and the brightness of UDs in this region are weak ($0.7 I_{ph}$). Most penumbral grains dies away quickly after they enter this umbral region, which is an indication that the convection is inhibited by strong magnetic fields.

4.6 UC Intensity Observation

In Figure 4.6, the minimum intensity of UC is plotted against the observing time. The thick solid curve is the smoothed data, and the thin curve is for the original subsonic filtered data. According to the measurements in the Figure 2 in Kopp & Rabin (1992), the field strength corresponding to such 1.56 μm continuum intensity is around 2650 G. In that paper, the author also found that there exists a linear relation between the magnetic field and 1.56 μm continuum intensity, if the intensity is lower than $0.6 I_{ph}$. If this is verified, then the fluctuation of the field strength can also be seen from Figure 4.6.

The brightness temperature can also be derived from the continuum observation. Because the opacity reaches the minimum near 1.6 μm, the brightness temperature is the maximum, i.e., the deepest layer of the visible sphere of the sun. Due to Wilson effect (geometrical depression), the observation in a sunspot reaches even lower layer. According

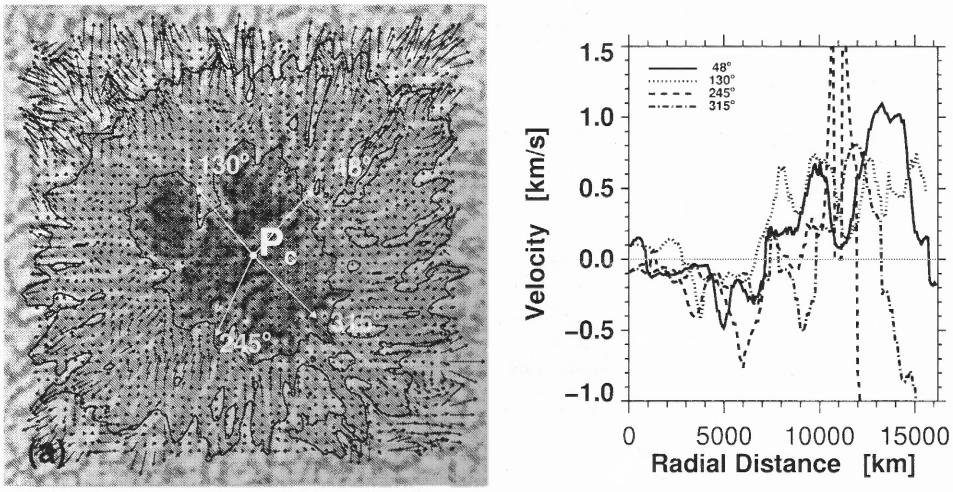

Figure 4.5 Proper motions along several directions. The flow calculated in these two figures are calculated by a simple LCT routine in IDL.

to the Planck's law of blackbody, the temperature and observed energy flux density (i.e., light intensity) is

$$T = \frac{hc}{\lambda k}\left[\ln\left(1+\frac{I_{ph}}{I}\{\exp\left(\frac{hc}{\lambda k T_{ph}}\right)-1\}\right)\right]^{-1} \tag{4.2}$$

4.7 Horizontal Velocity Map

A continuous time sequence of images can be used to derive the so-called proper motion, which represents the movements of features in the field of view. Erratic motion could be produced from the proper motion calculations. For example, the small and fine motions within an entity could be shadowed by the overall motion of that entity. It can be observed that the Evershed outflows embedded in the penumbral filamentary structures are moving with penumbral filaments. Therefore, the proper motion of the pixels within penumbral filaments do not reflect Evershed flows correctly. However, the proper motion calculation does show some insights of the flow pattern in a larger scale.

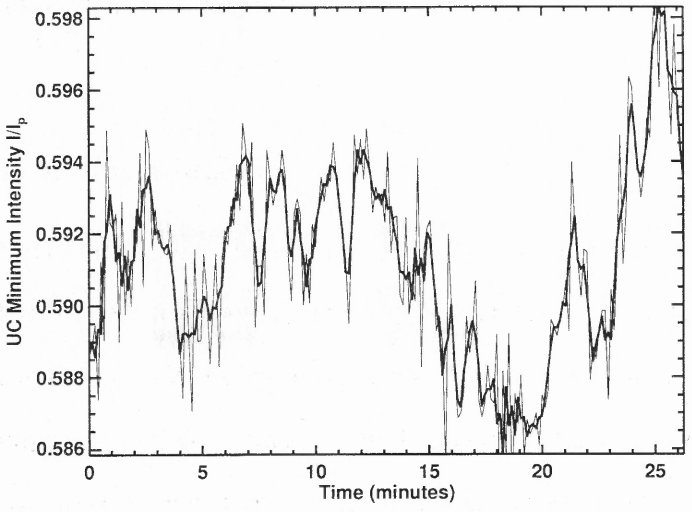

Figure 4.6 Minimum intensity plots of the umbral core. Data was processed with subsonic filter with $v_{ph} = 4$ km/s.

4.7.1 Calculation of Proper Motion

Proper motion is derived from the cross-correlation of pixel intensities of a sequence of images. It is also called "the optical flow" in many other literatures. A commonly used method in observational astronomy to calculate the proper motion was proposed by November & Simon (1988): "proper motion is defined as the displacement that maximizes the spatially localized cross correlation between two images of a scene separated by a sampling time delay τ that is smaller than the lifetime of tracers in the scene." In an equation form:

$$C(\delta,x) = \int J_t\left(\xi - \frac{\delta}{2}\right) J_{t+\tau}\left(\xi + \frac{\delta}{2}\right) W(x-\xi)d\xi \tag{4.3}$$

where, the window function $W(x)$ defines the spatial resolution of the proper motion calculation; $J(x)$ is the intensity map. This method is used to derive the velocity map in Figure 4.5 by applying OSLO package in SSW/IDL library.

4.7.2 Velocity Maps

In Figure 4.5, an 300-second averaged flow map is calculated from a set of subsonic-filtered (v_p=4 km/s Title et al. (1989)) images. In the overall picture, clear evidence of outflow pattern in penumbral region can be seen, along with an intermediate "region" where the flow pattern turns inward. The width of this transition region is not uniform even on average, see Figure 4.5. There is a 370 km wide transition region along certain direction — transition from outflow to inflow. By inspection of the Figure 4.5, it can be seen that the transition of the in/outflow of proper motions happens along the boundary between penumbral and its surroundings.

4.7.3 Flow Directions

The majority of the proper motion within the umbra is inflow, with small amplitudes on the order of only 100~200 m/s. This proves the findings by others (Wang & Zirin 1992, for example,) in other wavelengths — the UC region is a quiet region as regards to the proper motion. Since the Doppler map was not obtained for the same region, it is not possible to deduce the vertical motion of the flow. However, considering the continuity equation (November 1989):

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{v}) = 0 \tag{4.4}$$

where, $\mathbf{v}$ is the 3D velocity of plasma flow. It was shown that the vertical component $v_z = h_m (\nabla_h \cdot v_h)$, v_h is the planar velocity approximated by the proper motion; h_m is a scale height. So if the assumption that the flow is not shooting further up, but advect on the visible plane, the vertical flow is related to the divergence of the planar velocity field. Since the planar motion with UC is small, one shall expect to detect the vertical velocity field is also small in magnitude. The picture within the UC is, strong magnetic fields suppress the plasma flow. This is in good agrement with the magneto-convection model of sunspot in Rimmele (1997).

The situation within penumbra is complicated by Evershed flows. The apparent proper motion is in fact a superposition of the Evershed flow and the movements of the moving tubes (penumbral filaments). The movie shows that the filaments move inward, while the material of filaments moving outward. The superposition of the two motions can produce a dip in the velocity curve shown in Figure 4.5, for example, the curve for 48 °locating at around 12000 km.

4.8 Umbral Dots

Umbral dots (UDs) are isolated bright features embedded on the dark umbra background. The size of the dots were discovered ranging from $0''.1$ to $1''.4$. Note that large UD could be a cluster of unresolved small UDs. Therefore the real size of each individual UD is even smaller these numbers.

4.8.1 Filling Factor of Umbral Dots

The total area of UDs, identified by the region growing algorithm, as compared to the overall umbra region is also an interesting quantity. The ratio shows how the convective motion interact with strong magnetic field, and helps to prove the validity of popular sunspot models: cluster model (Parker 1979), monolithic flux tube (Meyer et al. 1974). In Figure 4.7, the ratio between the total area of UDs and the overall umbra area is plotted against the observation time. Discarding the uncertainty inherited from the algorithm identifying UDs, the ratio is oscillatory. The period is about 5 minutes. Is this just a coincidence with the 5-minute oscillation (Leighton et al. 1962; Ulrich 1970), or the dynamics of the UDs is correlated to the global oscillation? The answer is not available merely from the current observation.

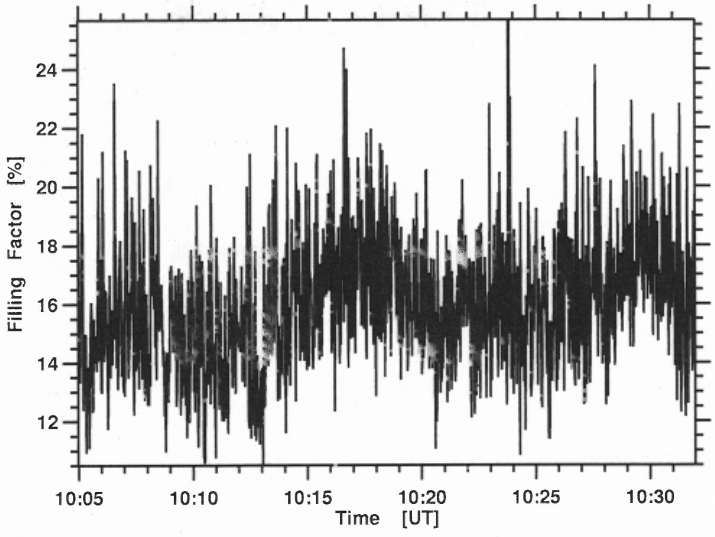

Figure 4.7 UD filling factor. This factor is the ratio between the total area of UDs and the corresponding overall area of umbra.

4.8.2 Number of Umbral Dots

The total number of UDs observed in 27 minutes oscillates, see Figure 4.8. A firm period for this oscillation was not observed. Since the seeing condition can introduce fake fractal structures in the field of view, one can compare the oscillation with the seeing condition variation shown in Figure 4.2 and did not find correlation between these two quantities. A possible source of the ambiguity comes from the intensity oscillation, such as 5-minute oscillation, because a constant threshold $0.8I_{ph}$ is introduced artificially for umbral region. However, the data set was processed with subsonic filter, $v_p = 4$ km/s. The photosphere oscillation should not play an obvious role in this scenario. On the other hand, the averaged number of UDs is 82.

4.8.3 Morphology of Umbral Dots

UDs observed in NOAA10707 can be classified into two categories by considering the pattern of formations: UD links and congregated groups. The later are confined by the former formations. There are at least two UD links can be confidently identified tightly

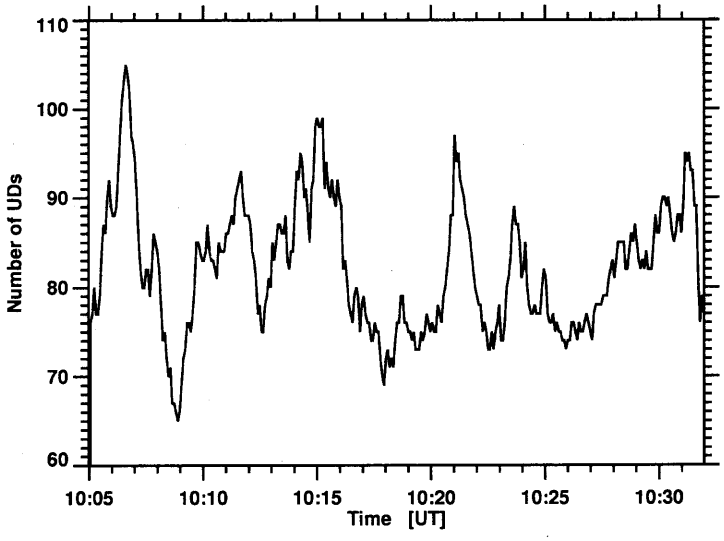

Figure 4.8 Variation of the number of UDs during the observation period.

connected, one even with a violent curvature. Probably another one is forming by four UDs with rather larger distances in between (about 0.25″).

Observed UDs are all have a brighter center surrounded by a fading boundary which is either connected to the fading boundaries of other UDs or outline itself from the local darker background. The observed shape of UDs are under the influences of instrument and seeing condition. Therefore, the size of a single UD can not be determined accurately. A rough estimate of the averaged UD size in our observation is 0″.30 (see Figure 4.9), and almost all of UDs are of elongated (stretched) shapes. There is also an outstanding penumbral grains (later a UD) with size of about 0″.8 to 1″.0, moving fast away from penumbra, and stalling after entering umbra without experiencing much intensity decreasing ($0.87 \sim 0.95 I_{ph}$) like other survived grains do.

The mean intensity of UDs are found to be primarily limited within 0.62~0.7 I_{ph}, and the corresponding diameter is typically below 0.7″. Note that the cutoff of diameter at 0.3″is determined by the available resolution at 1.56 μm.

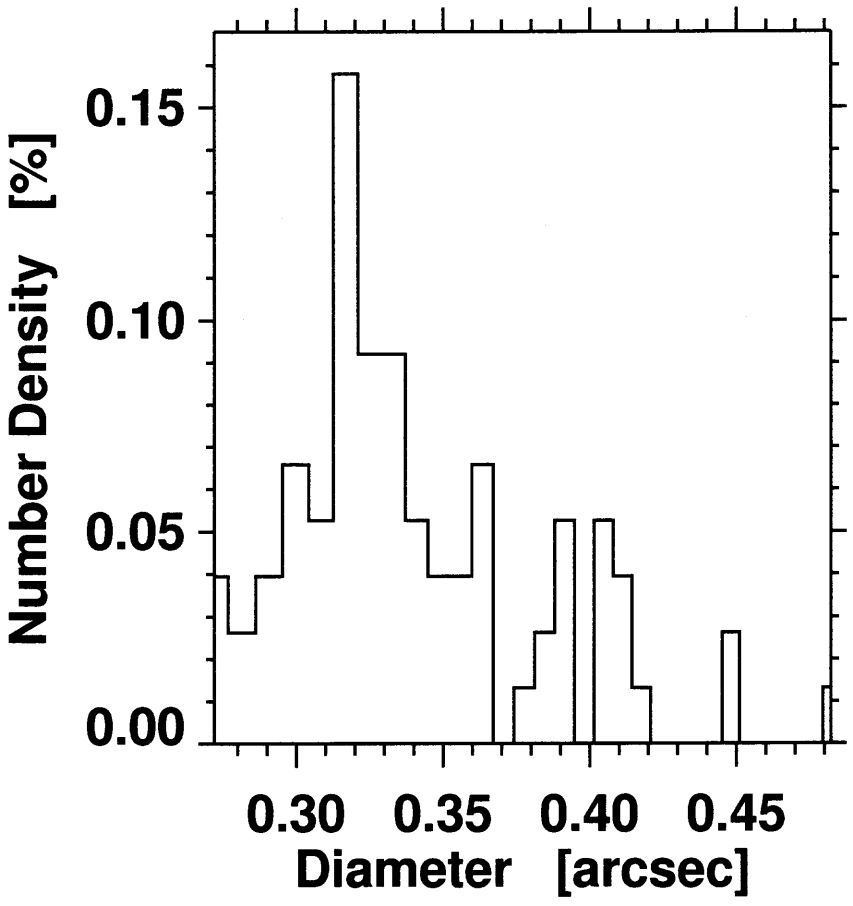

Figure 4.9 Umbral dots size histogram of the frame with highest rms-contrast.

4.9 Conclusions

1. In the decaying sunspot with averaged area 1.83×10^6 Mm^2, the averaged UD number is 82;

2. In such a sunspot, the area occupied by UDs is around 16% of the overall umbra region. This ratio oscillates with a period close to 5 minutes;

3. There exists a preferred size range and the corresponding mean intensity range for UDs: < 0.7 $''$in diameter and $0.62 \sim 0.7$ I_{ph};

4. The majority of UDs are of size of around $0.3''$in diameter according to the current spatial resolution;

5. The radial inflow dominates the penumbra and umbra regions. The typical inflow velocity in umbra is only $100 \sim 200$ m/s;

6. The is a transition belt between inflow and outflow, with a width of about 370 km in certain direction. The belt locates close to the outer boundary of penumbra;

7. The darkest region in umbra has minimum intensity 0.58 I_{ph} and varies with time.

CHAPTER 5

PHOTOSPHERIC MAGNETIC FIELDS IN CORONAL HOLE

5.1 Introduction

5.1.1 Small Scale Magnetic Fields

The small scale magnetic fields bear important clues of understanding the magnetic field evolution on the Sun apart from the large scale 11-year solar cycle. They occur predominantly in quiet sun regions, especially along the borders of supergranular cells, where the chromospheric network can be visualized in the filter grams of chromospheric lines. According to their locations and morphology, small scale magnetic fields can be classified in three categories, namely, network fields (Sánchez Almeida & Lites 2000; Wang et al. 1996), intranetwork (Lites & Socas-Navarro 2004), and ephemeral region (Harvey & Martin 1973).

In the first Stokes-V measurement using line ration technique by Keller et al. (1994), an upper limit on the intrinsic strength of intranetwork (IN) fields was set as 1000 G or 500 G, with 68% probability. The "typical" field strength of 500 G was found by Lin (1995) using infrared spectropolarimetry. Kneer & Stolpe (1996) presented an image in which small-scale magnetic elements possess substructure and are dynamical, with gas flows and magnetic field strength varying in space and time. Meunier et al. (1998) obtained the fraction of magnetic flux in a weak field form, i.e., with magnitude lower than 1000 G intrinsic strength in quiet sun. However, Socas-Navarro & Lites (2004); Socas-Navarro et al. (2004a,b) showed evidence of stronger ($\sim$1700 G) and weak ($<$500 G) fields coexisting within the resolution element in both network and IN regions, with a large fraction of area occupied by weak fields presenting convective upflows.

In the past, BBSO data were used to determine many important properties of small scale magnetic features in quiet sun regions. Wang et al. (1996) determined the flux distri-

bution in the regions, and the mean horizontal velocity fields of IN and network fields were determined by local correlation tracking; the lifetime of IN elements (Zhang et al. 1998a), the motion patterns and evolution of IN and network magnetic fields (Zhang et al. 1998b) were also investigated at BBSO.

5.1.2 Coronal Hole

Coronal holes are cool and low-density regions, which can be observed at both low-latitude and polar regions of the sun (Chiuderi Drago et al. 1999). They were first observed in X-ray by Underwood & Muney (1967), in EUV line by Reeves & Parkinson (1970), and in white light by Altschuler & Perry (1972). The magnetic fields within a coronal hole region are usually dominated by one polarity and the field lines in the upper atmosphere are open to the interplanetary region (Bohlin 1977), generating high-speed solar winds that can lead to geomagnetic storms (Krieger et al. 1973). According to the location and lifetime, there are three categories of coronal holes: polar, nonpoalr, and transient. Polar coronal holes have long lifetimes ($\sim$8 yr). Nonpolar coronal holes are usually associated with remnants of active regions and may persist for many solar rotations. Transient coronal holes are associated with eruptive events, such as filament eruptions and coronal mass ejections (CMEs) and have lifetimes of several days (Harvey & Recely 2002). Also, low-latitude coronal holes may show quasi-rigid rotation, and it has been suggested that magnetic reconnection must occur continuously at the boundary in order to maintain the integrity of the coronal holes (Kahler & Hudson 2002).

5.2 Comparison of The Fields in a Coronal Hole and a Quiet Region

5.2.1 Data Acquisition

The data used here was obtained in a joint observation of BBSO and the Transition Region and Coronal Explorer (TRACE) in the period of September 13-14 2004. The observation

achieved completed data sets on 14th and 16th, therefore, the data in these two days are of particular interests to the research of photospherical magnetic fields of coronal hole.

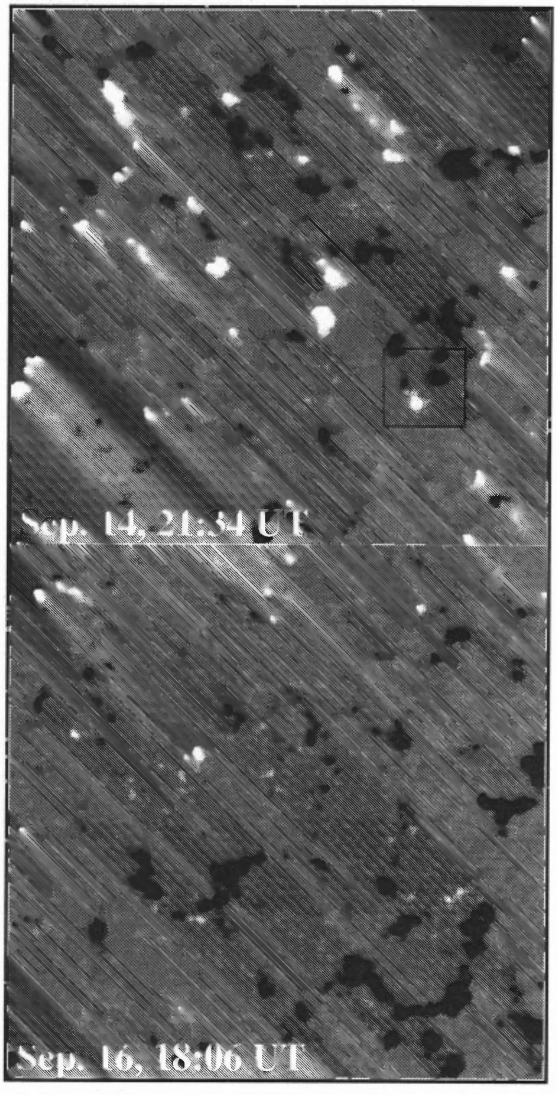

Figure 5.1 BBSO magnetograms in a quiet region (top) and a coronal hole (bottom). The field of view is $200'' \times 200''$. The box in the magnetogram of the quiet sun outlines a region of ephemeral flux (see Figure 5.2).

The data on 14th was a quiet region at N9°W13°; 16th data was a coronal hole at N31°E14°. The observation details are listed in Table 5.2.1. The observation in two days yields 60 magnetograms, 4 magnetograms per hour. Each magnetogram is an integration of 4096 frames spanning in about 15 minutes. Two samples of the magnetogram are shown

in Figure 5.1.

The evolution of the magnetic elements and corresponding dynamics are identified visually by investigating the movies at each wavelength. The first step is to find the ephemeral regions from the magnetogram and Hα movies visually. Second, the disappearance rate of magnetic flux in the two regions are studied according to the evolution of network elements. Third, the magnetic flux distribution is measured in the two areas.

5.2.2 Identification of Ephermal Active Regions

Ephemeral regions (ERs) can be identified as pairs of small, closely-spaced, opposite polarity features on the magnetograms, typically having a major axis of about 10,000 km (Martin & Harvey 1979). From the viewpoint of size only, ERs are hard to be discriminated from other small features such as network magnetic fields. The way to get around the misinterpretation is to follow the evolution of the magnetic features. ERs usually survive for only hours, and appear in pairs with similar magnitudes of flux.

Ephemeral regions are also seen in X-ray (Golub et al. 1981), Hα (Harvey & Martin 1973; De Pontieu 2002) and EUV (Roussel-Dupre et al. 1984). The X-ray bright points (XBP) can almost always be found associated with ERs, but not all ERs correspond to XBPs. The correspondence between ERs and Hα can also be used to ascertain the observed features.

Table 5.1 Setups of The Observations on 14th And 16th

Wavelength	Instrument	Observation
Hα	Koadak CCD, 25 cm refractor (BBSO)	90s cadence
CaII-K	same as above	same
H$\alpha \pm 0.60$Å	OSL CCD, 65 cm (BBSO)	30s cadence, FOV $210'' \times 210''$, image scale $0''.4$ pixel^{-1}
Magnetogram	Zeiss filter, DVMG, 25 cm refractor (BBSO)	$300'' \times 300''$, $0''.60$ pixel^{-1}
UV 1600Å	TRACE	Spatial $1''$, temporal 40s, FOV $250'' \times 250''$

In the current research, both Hα and CaI610.3 nm magnetogram movies are used to cross-check the features identified are ERs. On September 14 2004, 30 pairs of ERs in the quiet region are identified visually during a 7-hr observation. However the arch filaments in Hα only appear in 7 out of the 30 ERs. On September 16 2004 observation of coronal hole, only 17 pairs of bipolar ERs are found and one out of the 17 is connected by arch filament.

5.2.3 Magnetic Flux Emergence and Disappearance

The emergence rate of ERs were researched by various authors. In Hagenaar et al. (2003), based on six sequences of the full-disk MDI magnetograms from 1999 to 2001, the author found the emergence rate of ERs is 20 to 35×10^{-20} cm^{-2} day^{-1}. A similar result was obtained by Abramenko et al. (2006). However, the magnetogram date sets used in these observations are all from MDI data, which has detection limit of 17 G. The DVMG data at BBSO has much high sensitivity, namely 2 G, and the emerging rate found from 14th data is about 50×10^{-20} cm^{-2} day^{-1}. This discrepancy can be understood from the fact that BBSO data resolves more smaller ERs than MDI.

In the observation on 14th of quiet region, it can be clearly seen that the arch filament indicated by dotted line in Figure 5.2 connects the bipolar structure (22:49UT of H$\alpha-0.6$Åand 22:50UT of Hα dopplergram). At the same time, UV bright points corresponding the bipolar ER grow larger. In the end of the observation, the clusters of UV bright points scatters into many independent points without significantly losing the brightness.

The mean flux of an ER is 8.1×10^{18} Mx for quiet region, and 3.4×10^{18} Mx for the coronal hole region.

Figure 5.3 shows that, regarding the flux and flux density distribution, although there exist similarities between the quiet region and coronal region in general, ERs in coronal tend to possess less number of large flux and flux density. A natural consequence is that the total flux in the coronal hole region add up to 5.8×10^{19} Mx, as apposed to 2.5×10^{20} Mx

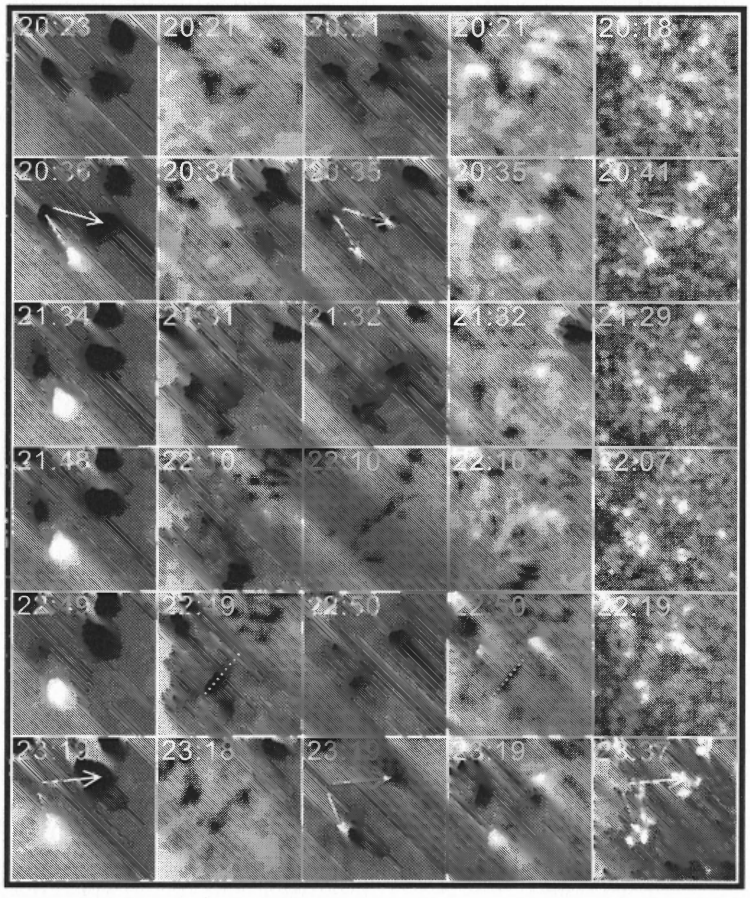

Figure 5.2 Quiet region, September 14th, 2004. From the left to the right, DVMG magnetogram, H$\alpha-0.6$Å, H$\alpha+0.6$Å, Hα Dopplergram, and UV1600ÅTRACE. The field of view is about $30''\times30''$. Arrows denote a bipolar ER; dotted line denote the location and orientation of an arch filament connecting the two polarities of the the bipolar ER. See Section 5.2.3 for the discussion.

in the quiet region. Therefore, the magnetic flux evolution in the coronal hole region is much slower than in the quiet region.

The arch filaments only appear in those ERs which have flux higher than 5.0×10^{18} Mx and flux density greater than 20 G. If arch filament structure of an ER is a closed magnetic loop, then the flux density of the loops in the observation ranges from 6 to 40 G in the quiet region, and 6 to 22 G in the coronal hole region.

The disappearance is mainly due to the cancellation of opposite polarity magnetic elements. The rate can be calculated according to the decrease rate of the total flux. In

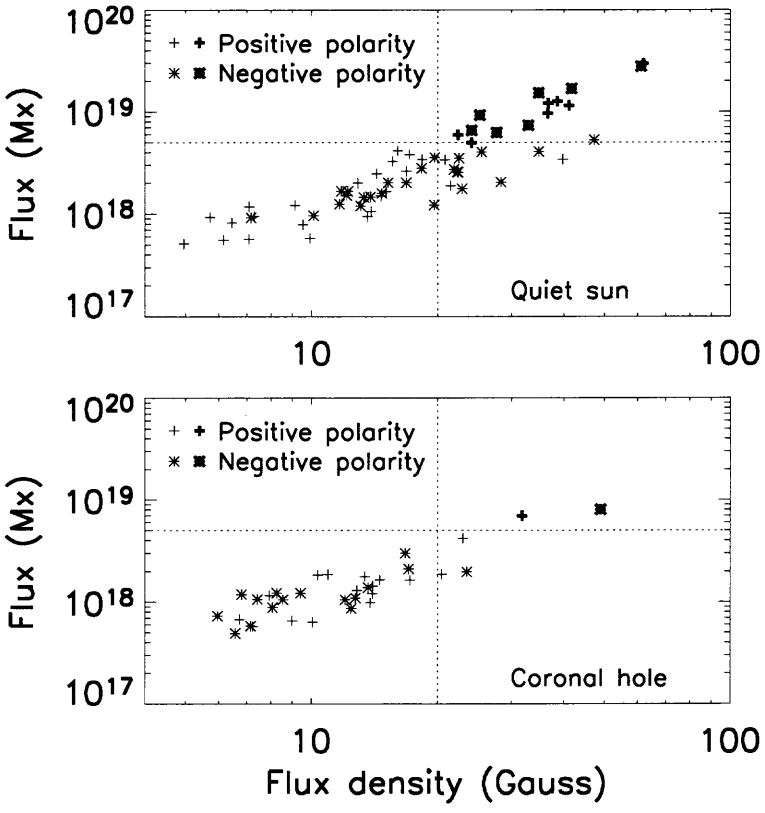

Figure 5.3 Flux vs flux density. Top: 30 pairs of ERs in the quiet region; bottom: 17 pairs of ERs in the coronal hole region. Vertical dotted lines stand for the location of flux density equals to 20 G; the horizontal dotted lines stand for the flux equal to 5×10^{18} Mx.

the quiet region, about 1.7×10^{20} Mx (2.1×10^{20} Mx) positive flux (negative flux) was lost during the 7 hr observation. In the coronal hole region, the numbers are 3.5×10^{19} Mx for positive polarity (9.6×10^{19} Mx for negative polarity).

5.2.4 Magnetic Flux Distribution

The magnetic flux distribution in quiet regions was presented by Wang et al. (1995); Socas-Navarro et al. (2004a); Domínguez Cerdeña et al. (2006). However, the distribution of IN flux in coronal holes is never revealed in details. Coronal holes lie in regions predominated by unipolar magnetic fields. Three criteria in Wang et al. (1995) are used to discriminate the IN elements and network elements. For each region, about 1000 magnetic elements

Table 5.2: Flux Distribution in Coronal Hole And Quiet Sun For Both Polarities

Parameters	Network (+)	Network (-)	IN (+)	IN (-)
		Quiet Sun		
Number	102	116	376	383
Total Flux (10^{20} Mx)	4.1	-6.4	2.2	-2.8
Flux imbalance	...	-0.22	...	-0.12
Mean flux (10^{18} Mx)	4.02	-5.52	0.59	-0.73
Flux density (G)	21.3	-25.8	4.1	-4.6
		Coronal Hole		
Number	59	153	540	342
Total Flux (10^{20} Mx)	1.1	-9.8	2.7	-1.6
Flux imbalance	...	-0.80	0.26	...
Mean flux (10^{18} Mx)	1.86	-6.41	0.50	-0.47
Flux density (G)	15.2	-29.5	4.2	-4.3

are identified and measured. The statistics are listed in Table 5.2. Figure 5.4 shows flux distributions of all measured magnetic elements. The bin size for IN field is 5.0×10^{16} Mx, and 5.0×10^{17} Mx for the network fields. In the quiet region, the flux distribution of IN and network elements is similar in both polarities.

The flux distribution in the coronal hole is different. For network fields, there are no positive elements with a magnetic flux about 4.0×10^{18} Mx. However, for the IN fields, the number of positive elements is much larger than the number of negative elements. This indicates that in the coronal hole, one polarity has both stronger flux and population, the other has weaker flux and less populated.

Another quantity can give insight of the flux distribution is

$$\Phi(B) = \int_S B ds \quad , \quad (B \geq B_t) \tag{5.1}$$

where B_t denotes the threshold flux density, and ds is the area element of integral. Applying this functional calculation to both the quiet region and the coronal hole region, the plots are shown in Figure 5.5. As the threshold flux density increases from 2 G (approximately the noise level) to 30 G, the flux in both polarities decreases uniformly in the quiet region. In the coronal hole, the positive flux decreases more quickly than the negative flux, with

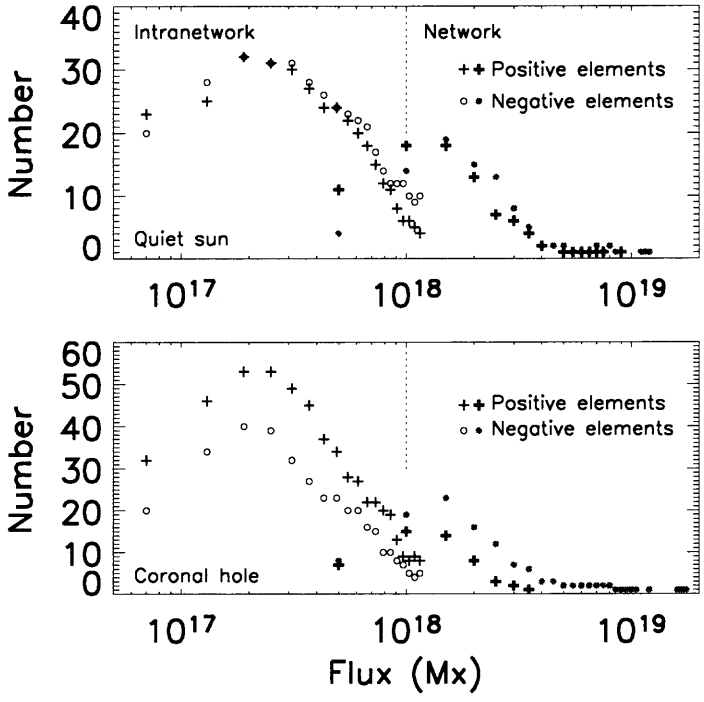

Figure 5.4 Flux distributions of positive and negative elements in the quiet region (top) and the coronal hole region (bottom). The dotted straight line represent the magnetic flux of 10^{18} Mx, which is chosen as the threshold to discriminate IN elements and network elements.

the increasing threshold flux density (middle panel of Figure 5.5). The ratio plot between the negative flux to the total flux shows that, for the coronal hole, 73% of the magnetic fields with a flux density greater than 2 G are negative, and 95% of the magnetic fields are negative if only the flux density greater than 20 G are integrated. Note that 20 G is the noise level of a typical magnetogram, non-BBSO-DVMG data. Compared with the quiet region with the coronal hole region, it can be found that the ratio changes more slowly in the quiet region.

The result confirms that in coronal hole, stronger fields with one polarity over populate weaker fields with the opposite polarity, and with a large margin.

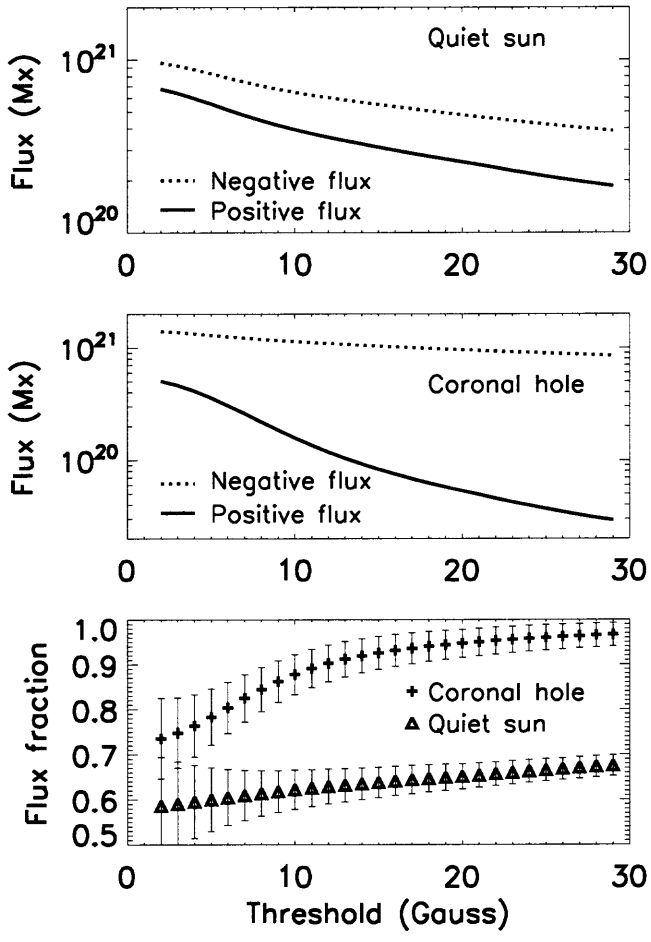

Figure 5.5 Variation of magnetic flux vs. threshold flux density in the field of view of magnetograms in Figure 5.1 in the quiet region (top) and the coronal hole (middle). The bottom panel plots the ratio between the negative flux and the total flux.

5.2.5 Is The Distribution An Isolated Case

The coronal hole data used so far comes only from one day observation. There could be a bias on an analysis of the distribution merely based on one day observation. However, the distribution shown in Figure 5.4 is also found in other data sets. For example, the VMG observation BBSO on June 4th 1992; DVMG observation of BBSO on September 17th 2004; and October 11th 2005.

5.2.6 Empirical Model of Coronal Hole

The emergence rate of the bipolar flux in the quiet region exceeds the rate in the coronal hole by a factor of 3. The scenario implied by this observation is in consistence with Fisk's model (Fisk 2005). Fisk's model indicates that coronal hole forms at locations where the rate of new flux emergence is locally minimum and open flux accumulates more easily. On the other hand, the flux disappearance rate in the quiet region is also twice as large as the rate in the coronal hole. The direct consequence of the discrepancy is that the magnetic energy converted to heat and kinetic energy is lowered in the coronal hole region.

The morphology of the magnetic structure in chromosphere is indicated by the Hα observation. Many of the observed Hα filamentary structures connect the bipolar features in the quiet region. In the coronal hole region, on the contrary, most of the observed Hα structures resemble jets, indicating locally open field lines.

Based on these observations, a schematic view of the 3D topology of the magnetic field can be assumed as shown in Figure 5.6. The loops in coronal holes are on average flatter than those in quiet regions. High and long closed loops are extremely rare, whereas short and low-lying loops are almost as abundant in coronal holes as in quiet regions (Wiegelmann & Solanki 2004).

The observations also hint that IN flux may be topologically connected to the network field (Zhang et al. 1999). It needs to be pointed out that IN flux having the same polarity as the surrounding network is more likely to be counted as network flux, and they are more rapidly destroyed either by merging with the surrounding network field with same polarity or by diffusing. Furthermore, IN flux of the opposite polarity rises in a more concentrated fashion, and is more easily detected.

5.2.7 Connection to Solar Wind

Tu et al. (2005) established that the fast solar wind starts blowing out of a coronal hole at 5000 to 20000 km above the photosphere in magnetic funnels. In the current data on

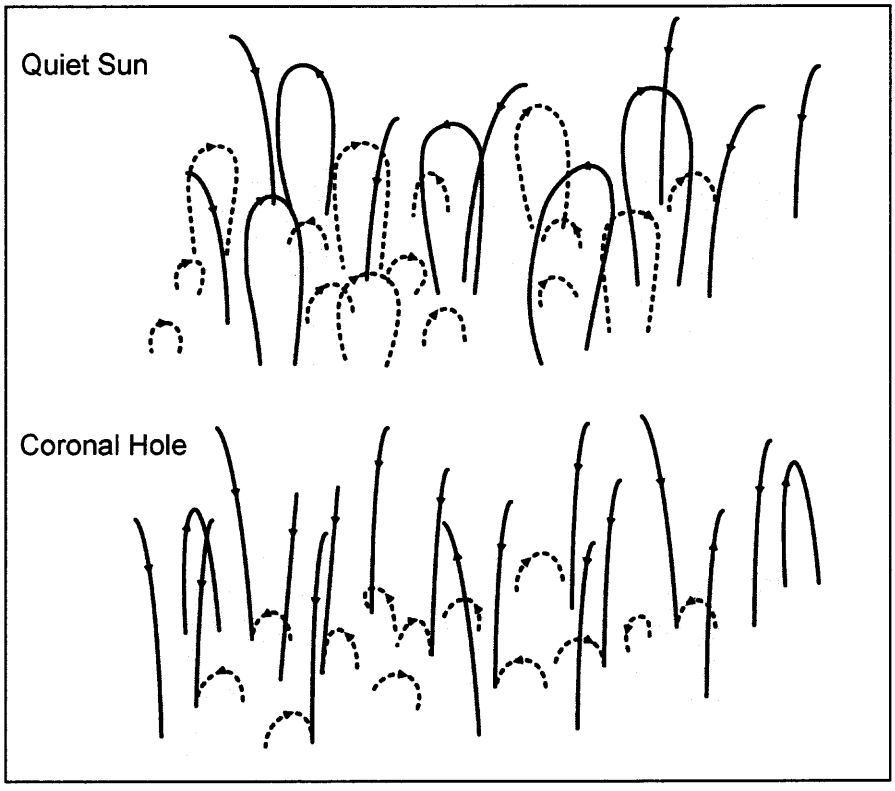

Figure 5.6 Schematic view of the magnetic structures in quiet region and coronal hole region. Solid curves stand for the magnetic field lines that have Hα counterparts; the closed lines show arch filaments, and the open-ended lines stand for macrospicules. The dotted lines indicate the magnetic field structures which have no Hα counterparts.

September 16th 2004, most closed magnetic loops were lower than 5000 km, and open magnetic fields with one polarity filled the space about 5000 km. Therefore, flux cancellation, or a lower magnetic reconnection in the photosphere and lower chromosphere, may only take place below the atmospheric level of 5000 km, not in the location where the fast solar wind originates. So the release of magnetic energy and the generation of fast solar wind take place at different levels.

5.3 Conclusions

Based on the accurate measurement of the magnetogram and other synchronized observational data, the following conclusion can be drawn about the small-scale magnetic fields

in the coronal hole region (Sep 14th 2004) as opposed to the quiet sun region (Sep 16th 2004):

1. The evolution of magnetic flux in the quiet region is much faster than that in the coronal hole region, as the flux appearing in the form of ephemeral regions in the quiet region is 4.3 times as fast as in the coronal hole region;
2. More magnetic elements with opposite polarities in the quiet region are connected by arch filaments, estimating from magnetograms and Hα images;
3. According to the measurement of roughly 1000 magnetic elements in both observed regions, for network fields in the coronal hole, there are many more negative elements than positive elements; for IN fields, there are many more positive elements than negative elements;
4. In the coronal hole region, the fraction of negative flux obviously changes with a different threshold flux density: About 73% of the magnetic fluxes with a flux density larger than 2 G have negative polarity, and the number increases to 95% if only those fields with flux density greater than 20 G are accounted;
5. In the coronal hole region, stronger fields are occupied by one predominant polarity; the majority of weaker fields are occupied by the opposite polarity.

For small scale magnetic fields observation, near infrared spectrum is a more suitable wavelength range to carry out the observation, due to larger Zeeman splitting and therefore higher sensitivity, lower scattering light level. Also, other more sophisticated methods, such as line ratio (Rabin 1992; Lin 1995) method, can be applied to the near infrared lines, for examples, Fe I λ1564.8 nm (g=3) and Fe I λ1565.3 nm (g=1.53) (Rueedi et al. 1995; Solanki et al. 1992, 1990). This task will be one of the most important goals of the future instruments in solar physics, such as NIRTF and IRIM.

CHAPTER 6

SUMMARY

In this thesis the concept design of a tunable filter system, NIRTF, was developed as part of the focal-plane instruments for ATST and NST (Chapter 2). The objective of the wavelength coverage is from 1 μm to 1.8 μm. NIRTF will be a unique and versatile instrument in future solar observations. Specifically, with an additional polarization analyzer, more accurate 2D magnetic flux density map can be deducted by the measurement of the line profiles at multiple spectral lines (Lin 1995).

The improvement of a similar system, IRIM/BBSO, was also discussed in the thesis (Section 3.5.2). The achromatic issue was not a major part of the original design of IRIM. This statement specifically refers to Lyot filter used in IRIM, and its polarization analyzer. The lack of achromatism of the filter system is primarily due to the waveplates used and the wavelength coverage of IRIM is limited within a few nanometers around 1564.85 nm. Therefore, the achromatic waveplate design (Chapter 3) is of critical importance in extending the capable functionality of IRIM.

The observation results using the Lyot filter in IRIM system are also presented in the thesis. Due to the opacity reaches minimum at 1.6 μm (Stix 2002), the obtained images of the sunspot (NOAA 10707) reveals the deepest layer on the photosphere. Fine structures within the sunspot are analyzed, specifically the horizontal flow pattern and umbral dots (Chapter 4).

The magnetic flux density measurement is the next step in the development of IRIM. Due to minimum scattering light (both in solar atmosphere and instrumental) (Kopp & Rabin 1992) in near infrared as opposed to visible, IRIM can be setup to measure the magnetic field within small and weak features such as umbral dots in sunspot. These measurements can be used to justify the debates (Socas-Navarro et al. 2004a; Schmidt & Balthasar 1994; Wiehr & Degenhardt 1993) regarding the physical nature of these fine structures within

sunspots.

Since the infrared lines has definite advantages in observing weak fields (Rueedi et al. 1995), IRIM (or NIRTF in the future) can provide unique observations on small scale weak magnetic features in quiet sun and coronal hole, both of which bear important signatures of the unrevealed underlying mechanisms producing weak magnetic flux on the sun.

As the first step, the historical DVMG observations of weak magnetic features on photosphere in quiet sun and a coronal hole are analyzed and very interesting distribution rules are found for the magnetic flux in these two regions (Chapter 5). And a 3D topological model is also proposed for the magnetic field structures in these regions. Future observations using IRIM or NIRTF will measure the magnetic flux density instead of only flux, which can be used to justify the proposed model, and might provide more information regarding the driving mechanisms in solar plasma.

REFERENCES

Abdusamatov, H. I., & Krat, V. A. 1969, Sol. Phys., 9, 420.

Abramenko, V. I., Fisk, L. A., & Yurchyshyn, V. B. 2006, Astrophys. J. Lett., 641, L65.

Altschuler, M. D., & Perry, R. M. 1972, Sol. Phys., 23, 410.

Beckers, J. M.,Dickson, L., & Joyce, R. S. 1975, Appl. Opt., 14, 2061.

Beckers, J. M. 1971, Appl. Opt., 10, 973.

Beckers, J. M. 1972, Appl. Opt., 11, 681.

Bohlin, J. D. 1977, Sol. Phys., 51, 377.

Born, M., & Wolf, E. 1999, Principles of Optics : Electromagnetic Theory of Propagation, Interference and Diffraction of Light / Max Born and Emil Wolf, 1999.

Bruls, J. H. M., Lites, B. W., & Murphy, G. A. 1991, Solar Polarimetry, ed. L. J. November (Sunspot: NSO), 444.

Bumba, V., & Suda, J. 1980, Bulletin of the Astronomical Institutes of Czechoslovakia, 31, 101.

Cao, W., Jing, J., Ma, J., Xu, Y., Wang, H., & Goode, P. R. 2006, PASP, 118, 838.

Chiuderi Drago, F., Landi, E., Fludra, A., & Kerdraon, A. 1999, Astron. Astrophys., 348, 261.

Cowling, T. G. 1975, Nature, 255, 189.

Cowling, T. G. 1981, Annu. Rev. Astron. Astrophys., 19, 115.

De Pontieu, B. 2002, Astrophys. J., 569, 474.

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.

Didkovsky, L. V., Denker, C., Goode, P. R., Wang, H., & Rimmele, T. R. 2003, Astronomische Nachrichten, 324, 297.

Didkovsky, L. V., Kuhn, J. R., & Goode, P. R. 2004, Proc. SPIE, 5171, 333.

Domínguez Cerdeña, I., Sánchez Almeida, J., & Kneer, F. 2006, Astrophys. J., 636, 496.

Gandorfer, A. M. 1999, ASSL Vol. 243: Polarization, 297.

Gary, G. A., West, E. A., Rees, D., McKay, J. A., Zukic, M., & Herman, P. 2007, Astron. Astrophys., 461, 707.

Evans, J. W. 1949, Journal of the Optical Society of America (1917-1983), 39, 229.

Filinski, I., & Skettrup, T. 1984, Appl. Opt., 23, 2747.

Fisk, L. A. 2005, Astrophys. J., 626, 563.

Flanders, D. C. 1983, Applied Physics Letters, 42, 492.

Geman, S. & Geman, G., IEEE Transactions, PAMI, 6, 721-741.

Ghosh, G. 1999, Optics Communications, 163, 95.

Golub, L., Rosner, R., Vaiana, G. S., & Weiss, N. O. 1981, Astrophys. J., 243, 309.

Guimond, T., & Elmore, D. 2004, SPIE's OE Magazine, V4, N5, P26.

Hagenaar, H. J., Schrijver, C. J., & Title, A. M. 2003, Astrophys. J., 584, 1107.

Hariharan, P. 1995, Measurement Science and Technology, 6, 1078.

Harvey, K. L., & Martin, S. F. 1973, Sol. Phys., 32, 389.

Harvey, K. L., & Recely, F. 2002, Sol. Phys., 211, 31.

Howard, R., & Labonte, B. J. 1981, Sol. Phys., 74, 131.

Ives, D., & Bezawada, N. 2007, Nuclear Instruments and Methods in Physics Research A, 573, 107.

Jones, R. C. 1941, Journal of the Optical Society of America, 31, 488.

Jones, R. C. 1941, Journal of the Optical Society of America, 31, 500.

Jones, R. C. 1942, Journal of the Optical Society of America, 32, 486.

Jones, R. C. 1947, Journal of the Optical Society of America, 37, 107.

Kahler, S. W., & Hudson, H. S. 2002, Astrophys. J., 574, 467.

Keil, S., et al. 2004, Proc. SPIE, 5489, 625.

Keller, C. U., Deubner, F.-L., Egger, U., Fleck, B., & Povel, H. P. 1994, Astron. Astrophys., 286, 626.

Kikuta, H., Ohira, Y., & Iwata, K. 1997, Appl. Opt., 36, 1566.

Kirkpatrick, S., Gelatt, C. D., & Vecchi, M. P. 1983, Science, 220, 671.

Kitai, R. 1986, Sol. Phys., 104, 287.

Kneer, F., & Stolpe, F. 1996, Sol. Phys., 164, 303.

Knobloch, E., & Weiss, N. O. 1983, Physica D Nonlinear Phenomena, 9, 379.

Kopp, G., & Rabin, D. 1992, Sol. Phys., 141, 253.

Krieger, A. S., Timothy, A. F., & Roelof, E. C. 1973, Sol. Phys., 29, 505.

Kuhn, J. R., Balasubramaniam, K. S., Kopp, G., Penn, M. J., Dombard, A. J., & Lin, H. 1994, Sol. Phys., 153, 143.

Leighton, R. B., Noyes, R. W., & Simon, G. W. 1962, Astrophys. J., 135, 474.

Lin, H. 1995, Astrophys. J., 446, 421.

Lites, B. W., Bida, T. A., Johannesson, A., & Scharmer, G. B. 1991, Astrophys. J., 373, 683.

Lites, B. W., & Socas-Navarro, H. 2004, Astrophys. J., 613, 600.

Loughhead, R. E. 1974, Sol. Phys., 38, 77.

Ma, J., Wang, J., Cao, W., Denker, C., & Wang, H. 2004, Proc. SPIE, 5523, 139.

Maltby, P., Avrett, E. H., Carlsson, M., Kjeldseth-Moe, O., Kurucz, R. L., & Loeser, R. 1986, Astrophys. J., 306, 284.

Martin, S. F., & Harvey, K. H. 1979, Sol. Phys., 64, 93.

McIntyre, C. M., & Harris, S. E. 1968, Journal of the Optical Society of America (1917-1983), 58, 1575.

Meunier, N., Solanki, S. K., & Livingston, W. C. 1998, Astron. Astrophys., 331, 771.

Meyer, F., Schmidt, H. U., Wilson, P. R., & Weiss, N. O. 1974, Mon. Not. R. Astron. Soc., 169, 35.

Meyer, F., Schmidt, H. U., & Weiss, N. O. 1977, Mon. Not. R. Astron. Soc., 179, 741.

Mickey, D. L., Canfield, R. C., Labonte, B. J., Leka, K. D., Waterson, M. F., & Weber, H. M. 1996, Sol. Phys., 168, 229.

Mogulskii, A. A., 1974, Theor. Probability Appl. 19 (726-736).

November, L. J. 1989, Astrophys. J., 344, 494.

November, L. J., & Simon, G. W. 1988, Astrophys. J., 333, 427.

Okunev, O. V., Domínguez Cerdeña, I., Puschmann, K. G., Kneer, F., & Sánchez Almeida, J. 2005, Astronomische Nachrichten, 326, 205.

Otten, M. & van GinnekenL, P., 1989, The Simulated Annealing Algorithm, Boston : Kluwer Academic Publishers, 1989.

Pancharatnam, S. 1955, Proc. Indian. Acad. Sci. (137-144), 41.

Parker, E. N. 1974, Sol. Phys., 36, 249.

Parker, E. N. 1979, Astrophys. J., 230, 905.

Petrovay, K., & Szakaly, G. 1993, Astron. Astrophys., 274, 543.

Rabin, D. 1992, Astrophys. J. Lett., 390, L103.

Reeves, E. M., & Parkinson, W. H. 1970, Astrophys. J. Suppl. Ser., 21, 1.

Ren, D., Hegwer, S. L., Rimmele, T., Didkovsky, L. V., & Goode, P. R. 2003, Proc. SPIE, 4853, 593.

Rimmele, T. R. 1997, Astrophys. J., 490, 458.

Rimmele, T. R. 2000, Proc. SPIE, 4007, 218.

Rimmele, T. R. 2004, Astrophys. J., 604, 906.

Rimmele, T. R., et al. 2003, Proc. SPIE, 4839, 635.

Rimmele, T. R., et al. 2004, Proc. SPIE, 5492, 944.

Rimmele, T. R., et al. 2005, Proc. SPIE, 5901, 41.

Rimmele, T., & Marino, J. 2006, Astrophys. J., 646, 593.

Roussel-Dupre, R., Wrathall, J., Nicolas, K. R., Bartoe, J. D. F., & Brueckner, G. E. 1984, Astrophys. J., 278, 428.

Rueedi, I., Solanki, S. K., Livingston, W., & Harvey, J. 1995, Astron. Astrophys. Suppl., 113, 91.

Sánchez Almeida, J., & Lites, B. W. 2000, Astrophys. J., 532, 1215.

Schilder, M. 1966, Trans. Am. Math. Soc. 125, (63-85).

Schmidt, W., & Balthasar, H. 1994, Astron. Astrophys., 283, 241.

Schüssler, M., & Baumann, I. 2006, Astron. Astrophys., 459, 945.

Schüssler, M., Vögler, A. 2006, Astrophys. J. Lett., 641, L73.

Severnyi, A. B. 1965, Soviet Astronomy, 9, 171.

Shurcliff, W. A. 1966, Cambridge, Mass.: Harvard University Press, 1966.

Sobotka, M., Bonet, J. A., & Vazquez, M. 1992, Astron. Astrophys., 257, 757.

Sobotka, M., Bonet, J. A., & Vazquez, M. 1992, Astron. Astrophys., 260, 437.

Sobotka, M., Bonet, J. A., & Vazquez, M. 1994, Astrophys. J., 426, 404.

Sobotka, M., Brandt, P. N., & Simon, G. W. 1997, Astron. Astrophys., 328, 682.

Sobotka, M., Brandt, P. N., & Simon, G. W. 1997, Astron. Astrophys., 328, 689.

Socas-Navarro, H. 2001, ASP Conf. Ser. 236: Advanced Solar Polarimetry – Theory, Observation, and Instrumentation, 236, 487.

Socas-Navarro, H., & Lites, B. W. 2004, Astrophys. J., 616, 587.

Socas-Navarro, H., Martínez Pillet, V., & Lites, B. W. 2004, Astrophys. J., 611, 1139.

Socas-Navarro, H., Pillet, V. M., Sobotka, M., & Vázquez, M. 2004, Astrophys. J., 614, 448.

Socas-Navarro, H., Trujillo Bueno, J., & Ruiz Cobo, B. 2000, Astrophys. J., 530, 977.

Solanki, S. K., Biemont, E., & Muerset, U. 1990, Astron. Astrophys. Suppl., 83, 307.

Solanki, S. K., Rueedi, I. K., & Livingston, W. 1992, Astron. Astrophys., 263, 312.

Solanki, S. K., Schüssler, M., & Fligge, M. 2000, Nature, 408, 445.

Solanki, S. K. 2003, Astron. Astrophys. Rev., 11, 153.

Solanki, S. K., Inhester, B., & Schüssler, M. 2006, Reports of Progress in Physics, 69, 563.

Spirock, T. J. 2005, Ph.D. Thesis, Big Bear Solar Observatory & New Jersey Institute of Technology.

Stix, M. 2002, The Sun : An Introduction – 2nd ed. /Michael Stix. Berlin : Springer, 2002. QB 521, S75.

Strous, L. 1994, Utrecht University, Ph.D. Thesis.

Thomas, J. H., Weiss, N. O., Tobias, S. M., & Brummell, N. H. 2002, Nature, 420, 390.

Title, A. M. 1975, Appl. Opt., 14, 229.

Title, A. M., Tarbell, T. D., Topka, K. P., Ferguson, S. H., Shine, R. A., & SOUP Team 1989, Astrophys. J., 336, 475.

Tritschler, A., & Schmidt, W. 1997, Astron. Astrophys., 321, 643.

Tritschler, A., Schmidt, W., Langhans, K., & Kentischer, T. 2002, Sol. Phys., 211, 17.

Tu, C.-Y., Zhou, C., Marsch, E., Xia, L.-D., Zhao, L., Wang, J.-X., & Wilhelm, K. 2005, Science, 308, 519.

Ulrich, R. K. 1970, Astrophys. J., 162, 993.

Underwood, J. H., & Muney, W. S. 1967, Sol. Phys., 1, 129.

Unno, W. 1956, PASJ, 8, 108.

Vernazza, J. E., Avrett, E. H., & Loeser, R. 1976, Astrophys. J. Suppl. Ser., 30, 1.

Volkmer, R., et al. 2006, Proc. SPIE, 6267.

Wang, H., Tang, F., Zirin, H., & Wang, J. 1996, Sol. Phys., 165, 223.

Wang, H., & Zirin, H. 1992, Sol. Phys., 140, 41.

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.

Wang, J., Wang, H., Tang, F., Lee, J. W., & Zirin, H. 1995, Sol. Phys., 160, 277.

Wang, J., & Ma, J. 2004, NIRTF Critical Design, ATST Technical Report.

Weiss, N. O., Brownjohn, D. P., Matthews, P. C., & Proctor, M. R. E. 1996, Mon. Not. R. Astron. Soc., 283, 1153.

Wiegelmann, T., & Solanki, S. K. 2004, Sol. Phys., 225, 227.

Wiehr, E., & Degenhardt, D. 1993, Astron. Astrophys., 278, 584.

Wiehr, E. 1994, Astron. Astrophys., 287, L1.

Ye, C. 2004, Appl. Opt., 43, 4007.

Zhang, J., Lin, G., Wang, J., Wang, H., & Zirin, H. 1998, Astron. Astrophys., 338, 322.

Zhang, J., Lin, G., Wang, J., Wang, H., & Zirin, H. 1998, Sol. Phys., 178, 245.

Zhang, J., Wang, J., Deng, Y., & Wang, H. 1999, Sol. Phys., 188, 47.

Zirin, H. 1974, Sol. Phys., 38, 91.