DISSERTATION

THE FARADAY FILTER-BASED SPECTROMETER: AN INSTRUMENT TO STUDY SODIUM NIGHTGLOW AND ASSOCIATED SODIUM AND OXYGEN CHEMISTRY IN THE MESOPAUSE REGION

Submitted by

Sean David Harrell

Department of Physics

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Summer 2010

COLORADO STATE UNIVERSITY

June 21, 2010

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY SEAN DAVID HARRELL ENTITLED THE FARADAY FILTER-BASED SPECTROMETER: AN INSTRUMENT TO STUDY SODIUM NIGHTGLOW AND ASSOCIATED SODIUM AND OXYGEN CHEMISTRY IN THE MESOPAUSE REGION BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate work

_____
Jacob L. Roberts

_____
Steven C. Reising

_____
Advisor: Chiao-Yao She

_____
Co-Advisor: David A. Krueger

_____
Department Chair: Hans Dieter Hochheimer

ABSTRACT OF DISSERTATION

THE FARADAY FILTER-BASED SPECTROMETER: AN INSTRUMENT TO STUDY SODIUM NIGHTGLOW AND ASSOCIATED SODIUM AND OXYGEN CHEMISTRY IN THE MESOPAUSE REGION

The newly developed Faraday Filter-Based Spectrometer (FFBS) makes possible spectroscopic study of the sodium nightglow in the mesopause region (80-110 km) of the atmosphere. This dissertation details the theory, design, and initial results of this instrument. The ratio of various combinations of NaD_2 and NaD_1 emission intensities can provide information on the oxygen and sodium chemistry in the mesosphere and lower thermosphere (MLT) region.

Early understanding of the production of sodium nightglow utilized the series of chemical reactions known as the Chapman Mechanism. This mechanism involves both sodium and various oxygen species to produce excited states of sodium, which then relax to the ground state and emit light. The emitted light is centered at two wavelengths: D_2 (589.158 nm) and D_1 (589.756 nm). If the excited states are populated according to the statistical weights of their spin-orbit coupling the $R_D = D_2/D_1$ intensity ratio should be 2; however there is no a priori reason that the spin-orbit states should be populated statistically in the Chapman mechanism. While early measurement of R_D yields a value

of 1.98±0.1, more recent measurements show a variation from 1.3 to 1.8; it peaks at the equinoxes and reaches minimum at the solstices. A possible explanation for this variation utilizes a modification to the Chapman Mechanism, which relates the R_D value to variations in the atomic oxygen to molecular oxygen, $[O]/[O_2]$, concentration ratio through two different chemical pathways for sodium nightglow production.

The FFBS is designed to measure R_D, the fractional contribution of the two chemical pathways of the modified Chapman Mechanism, and other parameters which are directly proportional to $[O]/[O_2]$. These parameters will help to test the validity of the modified Chapman mechanism. The delineation of the two chemical pathways requires an instrument with a spectral resolution of 0.0002 nm, something that is not possible with traditional spectroscopic instruments. The solution presented here utilizes two ultra narrowband sodium vapor Faraday filters. These utilize the Faraday rotation of light due to the Zeeman splitting of energy levels of sodium atoms in a vapor in an axial magnetic field between crossed polarizers to create an optical filter near both the D_1 and D_2 resonances with a full width, half maximum bandwidth of approximately 4 GHz (0.004 nm). This leads to a resolution that is good enough to distinguish the two different sodium nightglow spectral linewidths produced by the two pathways of the modified Chapman Mechanism which differ by 1.7 GHz. As a result, the FFBS is able to determine the fractional contribution from each pathway, as well as R_D. Therefore the FFBS provides a new method of investigating $[O]/[O_2]$ in the mesopause region remotely from the ground.

Data from this instrument supports the previous conclusion of a varying R_D, with 2009 autumnal equinox measurements averaging around 1.68 and 2010 vernal equinox

measurements averaging around 1.52, qualitatively in agreement with previous results by Slanger et al. (2005). In addition to R_D, the first known measurements of parameters specific to the modified Chapman mechanism are presented and discussed.

The dissertation concludes with a discussion of future work needed to convert the FFBS measurements into $[O]/[O_2]$ values, as well as future plans for the FFBS instrument.

Sean David Harrell
Department of Physics
Colorado State University
Fort Collins, CO 80523
Summer 2010

Acknowledgements

I would like to thank all of the people who have assisted me in the past six years to complete my dissertation. First, I would like to thank my advisor, Joe She, and my co-advisor, David Krueger. Their dedicated guidance, encouragement, and constant questioning have taught me the skills needed to be an excellent researcher. Specifically as Joe heads into retirement, I thank him for guiding me as his last PhD student.

I have been privileged to work with the lidar group at CSU. Members Tao Li, Phil Acott, Jia Yue, and Titus Yuan all helped with useful discussions, with the difficult task of 24-hour lidar observations, and with figuring out how my nightglow instrument was supposed to operate. Bob Adame in the department machine shop also never shied away from dusting off a little-used tool to assist me in making one more oddly-shaped part. Undergraduate students Jason Hahn, Stefan Schweiger, and Andrew Shelton also provided assistance in the construction of my spectrometer. Jason and Stefan deserve credit for dealing with the difficult process of figuring out how the magnet bars worked.

There are several outside collaborators whose help was greatly appreciated. Xinzhao Chu and her group at the University of Colorado at Boulder helped with several nights of observation. Jonathan Friedman, Wentao Huang, and Biff Williams provided many useful discussions about ways to improve Faraday filter design. Most importantly, discussions with John Plane and Tom Slanger were essential to my understanding of the theory of nightglow production.

Finally, and most importantly, I need to thank my wife, Jess. Without her continual support and love through long nights of data collection, missed deadlines, and broken promises, this dissertation would never have been completed. Thank you.

Table of Contents

CHAPTER 1 INTRODUCTION 1
1.1 Atmospheric structure 2
1.2 Mesosphere-lower thermosphere 6
1.3 Outline of dissertation 8

CHAPTER 2 MESOPAUSE REGION CHEMISTRY 10
2.1 General Composition and Chemistry 10
2.2 Oxygen measurements: conclusions and limitations 13
2.3 Importance of atomic oxygen 14
2.4 Chemistry of Mesospheric sodium 15

CHAPTER 3 NIGHTGLOW 18
3.1 History 18
3.2 Sodium Nightglow Production I: The Chapman Mechanism 19
3.3 Measurements of Sipler and Biondi 20
3.4 Sodium nightglow production II: Measurements of Slanger *et al.* and the modified Chapman Mechanism 21
3.5 Sodium nightglow production III: Chemical rate equations and implications for nightglow measurements 24
3.6 Sodium nightglow production IV: Reaction broadened nightglow spectra 27
3.6.1 Specific values for the modified Chapman Mechanism 29
3.6.2 Discussion 30
3.7 Necessity of the Faraday Filter-Based Spectrometer 33

CHAPTER 4 FARADAY FILTER PHYSICS 34
4.1 Classical theory of filter transmission 36
4.2 Quantum-mechanical theory 39
4.2.1 Derivation of susceptibility 39
4.2.2 Derivation of Transmission matrix elements 42
4.3 Computer program results and discussion 47
4.4. Conclusions: Other atmospheric studies using Faraday filters 52

CHAPTER 5 THE FARADAY FILTER-BASED SPECTROMETER 54
5.1 FFBS optical layout 54
5.2 Data recording and electronic control 58
5.2.1 Stand-alone operation mode 58
5.2.2 Simultaneous lidar operation mode 60
5.3 Data analysis 61
5.3.1 Reduction of FFBS signals 61

5.3.2 Error propagation 65
5.4 Faraday filter design and temperature control 68
5.5 Scanning the Faraday and Interference filters 70
5.6 Measuring the γ passive transmission factor 74
5.7 Determining Faraday Filter temperature settings 75

CHAPTER 6 COLORADO STATE UNIVERSITY SODIUM LIDAR SYSTEM 78
6.1 Introduction to the CSU Sodium Lidar System 78
6.2 Lidar Basics 79
6.3 The lidar transmitter 79
6.3.1 Doppler free spectroscopy for laser locking 80
6.3.2 Transmitter conclusions 84
6.4 The electronic control system 85
6.6 Lidar results and the FFBS 87

CHAPTER 7 SPECTROMETER DATA SETS AND ANALYSIS 90
7.1 Data collection specifics 90
7.2 Data simulation 91
7.3 Faraday filter stability 96
7.4 Data Sets 97
7.4.1 Moonless, cloudless data sets 97
7.4.2 Data with lunar background 107
7.4.3 April 2010 data with different Faraday filter temperature settings 112
7.4.4 Variation of γ and technical issues 114
7.5 Conclusions 115

CHAPTER 8 CONCLUSIONS AND FUTURE WORK 117
8.1 Future Work 117
8.1.1 Instrument Improvement 118
8.1.2 Science Studies 119
8.2 Final Conclusions 122

Bibliography 123

APPENDIX A DETAILS OF THE QUANTUM-MECHANICAL CALCULATION OF ELECTRONIC TRANSITIONS 130
A.1 Derivation of transition matrix elements and relation to known parameters 130
A.2 D_1 transition probabilities of allowed transitions 136
A.3 D_2 transition probabilities of allowed transitions 137

APPENDIX B DETAILS OF THE CSU LIDAR SYSTEM 140
B.1 Spectral variation due to temperature and wind 140
B.2 The lidar equation 141
B.3 Acousto-optic modulation for frequency shifting 145
B.4 Pulse amplification and frequency chirp measurement 148
B.5 The lidar receiver 151

B.5.1 The telescopes and fiber coupling 152
B.5.2 The Faraday filter for lidar use 152
B.5.3 Photon Detection 154

APPENDIX C CURVE FITTING FOR DATA ANALYSIS 155
C.1 Curve fit procedure 155
C.2 Results from 26 September 2009 156
C.3 Discussion 157

List of Figures

Figure 1.1: Schematic of atmospheric thermal variation as a function of both altitude and pressure, with layer designations. From www.learner.org. 3

Figure 1.2: Plot of mean zonal wind in winter (left) and summer (right). The bottom of the plot shows the Gaussian distribution of gravity wave phase speeds; the shaded portion indicates the phase speeds that are filtered by the mean wind and therefore cannot transmission to the MLT. Those waves that do survive to the MLT break, and impart wave drag in the direction indicated. (from Lindzen 1981)..... 5

Figure 1.3 MLT region forcing. Left: the Coriolis force balances the gravity wave drag, or body force, to create a meridional flow from the summer to the winter pole in the MLT region (from Yue, 2009). Right: the meridional flow which leads to air rising in the summer, and falling in the winter, cooling the summer and warming the winter mesopause (from Holton and Alexander, 2000). 6

Figure 2.1: Mixing ratio profiles of hydrogen-containing compounds (left) and oxygen containing compounds (right) as calculated by the ROSE model. Hydrogen profiles are scaled by the number of hydrogen atoms they contain. Solid lines are for daytime, dashed lines for nighttime. From Smith (2004). 12

Figure 2.2: Logarithm of atomic oxygen mixing ratio as a function of latitude and height at four local times as calculated by the ROSE model. From Smith (2004). 13

Figure 2.3: Schematic showing significant chemical cycles of sodium in the mesopause region.Input of sodium comes from meteoric ablation, and sodium is removed from the system by the formation of meteor smoke, which is metallic compounds polymerizing together with silicon oxides, which condense and coagulate into particles which are too large to remain suspended above 80 km. From Plane (2003). 17

Figure 3.1: Nightglow results from Sipler and Biondi, taken with a scanning Fabry-Perot Interferometer. The average value of their nightglow measurements is 1.98±0.1. The times to the left and right of the dashed lines (local solar zenith angle > 100°) is during twilight. Twilight airglow becomes dominated by resonant florescence by sodium of solar photons; since the D_2 is 90% of the D_1 emission from the sun, coupled with the stronger D_2 absorption of sunlight at D_2 frequencies as the light traverses the sodium layer to reach the observation point, the observed R_D quickly decreases (Sipler and Biondi 1977). 21

Figure 3.2: Results from Slanger *et al*, showing annual variation in R_D (open points) and nightglow intensity (solid points) (left pane) and R_D vs. $[O]/[O_2]$ (right pane). For details, see Slanger et al. (2005). 22

Figure 3.3: Calculated normalized nightglow spectra for (A) D_1 and (B) D_2. The dotted curve in each figure represents the spectra produced by (R2); the dashed curve is that from reaction (R2b). This curve is for the case of one vibrational quantum of O_2 for both (R2) and (R2b). Also plotted in solid for comparison are the normalized spectra as predicted by the quantum mechanical hyperfine structure. The temperature for each curve is 200 K. 31

Figure 3.4: Difference between D_1 spectra produced by assuming a NaO_2 Boltzmann distribution of 50 K (solid) and 200 K (dashed). The spectra shown are for the NaO(X) or (R2b) case. 32

Figure 4.1: Schematic of an atomic vapor Faraday filter, consisting of a vapor cell in an axial magnetic field between crossed polarizers. 34

Figure 4.2: Basic, Fine and Hyperfine structure for 3S and 3P states of the Na atom (not to scale). 43

Figure 4.3: Energy level diagram for Na showing the $3^2P_{3/2}$ excited states. (a) shows the single fine structure state, with a degeneracy of 16. (b) is the hyperfine splitting case with no external magnetic field. The $|F>$ eigenstate notation and degeneracies (in parentheses) are indicated to the right and the scale in GHz is to the left. (c) shows the exact solution for the Zeeman structure for an external magnetic field of 2000G. The states to be broken up into 4 closely spaced groups due to the different values for μ_I and μ_B. Each Zeeman state has a degeneracy of 1. The $|m_I\, m_J>$ eigenstate notations are listed to the right and the scale in GHz is to the left. The numerical values of the splitting are listed in Table 4.1 and Table 4.2. 46

Figure 4.4: D_2 (a) and D_1 (b) transmission as a fraction of input linear polarization (solid) and Faraday rotation in units of π (dashed) vs. frequency (GHz) for a Na vapor Faraday filter optimized for maximium transmission at the D_2 line center. (c) and (d) are the same for a D_1 optimized filter. Filter parameters are listed in Table 4.5. 51

Figure 4.5: χ' and χ'' curves for Na D_2 (a) and D_1 (b) lines split by the Zeeman effect due to the 1850 Gauss external magnetic field for the D_2 optimized filter. Solid gray curve is χ' for σ_-, gray dashed is χ'' for σ_-, black solid is χ' for σ_+, and black dashed is χ'' for σ_+. 52

Figure 5.1: Layout of the FFBS. Solid lines with arrowheads indicate the light path, while dotted lines indicate electronic connections. 55

Figure 5.2: Photograph of the FFBS, with approximate dimensions and components indicated. Not shown are the two 40% quantum efficiency PMT's, which are outside the enclosure to the left of Faraday filter B and above the top of the picture for Faraday filter A, as shown in this figure. The fiber from the telescope is also not shown; it enters the FFBS enclosure and the filter wheel from the bottom of the photograph. 56

Figure 5.3: Time-resolved FFBS data collected while the lidar was running. Each point represents one range bin 150 m in length, or 1 μs. As can be seen, only the first 2-3 time bins have contamination from the lidar signal. Note that the first range bin is around 1.6 km; this is because Fort Collins is 1570 m altitude. 61

Figure 5.4: Cross-section of the Faraday filter design. The light enters the Faraday filter from the left, and exits from the right, where "Polarizer goes here" indicates the location of the Glan-Thompson (second) polarizer. 70

Figure 5.5: Diagram for scanning of Faraday filters (not to scale). For interference filter scanning, photodiode 3 is placed directly after the filter wheel. Absolute frequency is monitored by the Doppler free spectroscopy unit, and relative frequency is monitored by the Fabry-Perot interferometer. Laser power is monitored by photodiode 2, while the transmission function is measured by photodiode 3. 72

Figure 5.6: Example of a Faraday filter scan file. The black Fabry-Perot curve has peaks at every 750 MHz, this can be scaled to absolute frequency using the Doppler free spectrum (in green), which has the D_{2a} Lamb dip near the minimum voltage. The blue curve is the laser normalization channel, which is used to correct for varying laser power over the course of a scan. 73

Figure 5.7: A sample analyzed average of several Faraday filter scans, corrected for absolute transmission. The frequency scale is in GHz. 74

Figure 5.8: The theoretical (red) and experimental (blue) Faraday filter transmission vs frequency offset from line center for the two Faraday filters. Top Left: Faraday filter A D_1, Top Right: Faraday filter A D_2, Lower Left: Faraday filter B D_1 Lower Right: Faraday filter B D_2. The experimental data are the average of several scans. 77

Figure 6.1: The CSU Na lidar transmitter system. Figure from Li (2005). 81

Figure 6.2: Comparison of the Doppler Free spectrum (solid) and the Doppler broadened spectrum (dashed). Note the lamb dips at the D_{2a} and D_{2b} peaks, and the crossover peak halfway in between. For details see text. From Arnold and She (2003). 82

Figure 6.3: D_1 Doppler free spectrum, showing the three Lamb dips at D_{1a} and D_{1b} and the three crossover peaks. From She and Yu (1995). 84

Figure 6.4: The electronics timing diagram for the CSU lidar system. For details, see the text. "Count-through" is split to trigger both the lidar and FFBS when the FFBS is run simultaneously with the lidar. 86

Figure 6.5: CSU lidar sodium density contour plot. Scale is m^{-3}. This data is for one telescope, and from the night of UT day 289 of 2009. 88

Figure 7.1: Long-term Faraday filter stability. Top Left: Faraday filter A D_2 transmission, Top Right: Faraday filter A D_1 transmission, Bottom Left: Faraday filter B D_2 transmission, Bottom Right: Faraday filter B D_1 transmission. In each plot, black is the

measurement from 25 September 2009, red is from 14 January 2010, blue is from 22 February 2010, and green is from 17 March 2010. Note that Faraday filter A was dismantled before the 17 March measurement. 96

Figure 7.2: Nightglow data from 26 September 2009. (A) Raw signal with dark counts removed. Red: Faraday filter A D_1 signals (corrected by γ), blue: Faraday filter B D_1 signals, green: Faraday filter A D_2 signals (corrected by γ), and black: Faraday filter B D_2 signals. (B) R_{DX} (red) and R_{DA} (blue) measurements. (C) R_{D1} (red) and R_{D2} (blue) measurements integrated for 30 minutes. (D) R_D measurements. (B) through (D) have been integrated for 30 minutes and error bars due to photon noise. 100

Figure 7.3: Same as figure 7.2, except for 30 September 2009. 101

Figure 7.4: Same as figure 7.2, except for 7 October 2009. 102

Figure 7.5: Same as figure 7.2, except for 15 March 2010. (E) and (F) are lidar measured sodium density contour plots for 30° off zenith to the north and 20° off zenith to the east, respectively. 103

Figure 7.6: Same as figure 7.5, except for 16 March 2010. 104

Figure 7.7: Plot of the nightly average R_D values from Table 7.4, as a function of the day of the year. Solid points are the data from Slanger et al (2005), and open points are the FFBS results. The error bars in both cases are equal to one standard deviation. 106

Figure 7.8: Lunar signal for 23 February 2010. Left panel: raw signal from the 532 nm interference filter for Faraday filter A (open points) and Faraday filter B (solid points). Right panel: Smoothed, zero-point subtracted, and normalized signals, points are the same as the left panel. 110

Figure 7.9: Comparison of nightglow signals with and without lunar signal subtraction for 23 February 2010. Note that there is no dark count subtraction. Solid points are with no lunar signal subtraction, open points are after lunar subtraction. The sets of curves are as follows from top to bottom: D_2 signal for Faraday filter B, D_1 signals for Faraday filter B, D_1 signals for Faraday filter A, and D_2 signals for Faraday filter A. 110

Figure 7.10: Analyzed results from 23 February 2010. (A) R_{D1} (red) and R_{D2} (blue). (B) R_{DX} (red) and R_{DA} (blue). (C) R_D for lunar signal subtracted data (red) and for no subtraction (blue). 111

Figure 7.11: Faraday filter A functions from April 2010. Left panel: D_1 transmission, Right panel: D_2 transmission. 112

Figure 7.12: Data from 19 April 2010. (A) Raw signal with dark counts removed. Red: Faraday filter A D_1 signals (corrected by γ), blue: Faraday filter B D_1 signals, green: Faraday filter A D_2 signals (corrected by γ), and black: Faraday filter B D_2 signals. (B) R_{DX} (red) and R_{DA} (blue) measurements. (C) R_{D1} (red) and R_{D2} (blue) measurements

integrated for 30 minutes. (D) R_D measurements. (B) through (D) have are integrated for 30 minutes and error bars due to photon noise. 113

Figure 7.13: Plot of all FFBS results for R_D. Red points are moonless and cloudless data with measured γ; Blue points are moonless and cloudless data with interpolated γ; the Green point is moonlight-subtracted results; Black points are April 2010 data with increased error due to the modified *c* coefficients. The error bars are the standard deviation of the data sets. 116

Figure B.1: The Na hyperfine D_2 spectra normalized to unit area. (a) Spectra at 150 K (solid), 200 K (dashed) and 250 K (dotted), and 0 m/s wind, showing effects of temperature broadening. (b) Spectra at 200 K, and at three radial velocities: 0 m/s (solid), +50 m/s (dotted) and -50 m/s (dashed). Also indicated are the three lidar probe frequencies. 141

Figure B.2: Calibration curves for relating R_W and R_T to temperature and LOS velocities. 144

Figure B. 3: The dual-pass acousto-optic modulator optical layout. Part (a) shows the path of the unshifted beam passing at the D_{2a} peak. (b) and (c) show the ±630 MHz shifted beams. 147

Figure B.4: The chirp subsystem. PD is photodiode, BS is beamsplitter. From Yuan et al. (2009). 150

Figure B.5: The chirp transmission and calibration curve. (A) Transmission function: the solid line is the I_2 spectrum when probed with the CW laser, the dashed line is the spectrum when probed with the PDA output pulses, and the dotted line shows the relative placement and width of the three-frequency PDA output pulses. (B) The polynomial fit to the chirp frequency shift (GHz) vs. chirp ratio R_c. 151

Figure B.6: Sample Faraday filter transmission for lidar. Notice that the transmission is scaled to unity at the peak for convinience; this is because the lidar data analysis only relies upon relative filter transmission. 153

Figure C.1: 26 September 2009 data analyzed with the linear curve fit method. A. Calculated nightglow intensities: Red: I_{1X}, Blue: $I_{1A,}$ Green: I_{2X}, Black: I_{2A}. B. R_{DX} (Red) and R_{DA} (Blue). C. R_{D1} (Red) and R_{D2} (Blue). D. R_D. Data from 0.5 hour intervals. Solid points are the fit before 6 UT, and open points are after 6 UT. Error bars are calculated by propagating the error in the fitting parameters. 158

List of Tables

Table 3.1: Na atom energies and velocities for different possible O_2 vibrational excitations and corresponding Boltzmann temperatures. 30

Table 4.1: Hyperfine structure state offset from fine structure (GHz). 46

Table 4.2: Zeeman effect state offset (GHz) from fine structure for B_0=2000 G. 47

Table 4.3: Necessary values for calculation. 50

Table 4.4: Linestrength, transition vacuum wavelength and linewidth independent of isotope 50

Table 4.5: Filter Parameters used to generate Figs. 4.4-4.5, and transition frequencies .. 50

Table 7.1: Results of the FFBS data simulation. For spectra "original" indicates the 21219 cm¬-1 NaO bond energy derived spectra, including one vibrational excitation for both A and X reaction pathways. Unless otherwise indicated, the temperature of the COM of the NaO + O system used is 200 K, and the line of sight wind is 0 m/s. "Sipler and Biondi" refers to using the upper and lower limits as found in their 1978 paper, namely a Na atom dissociation kinetic energy of 282 cm^{-1} for reaction R2, and 403 cm^{-1} for reaction R2b. The data were generated using R_D=1.7, R_{D2}=2.0, and R_{D1}=1.5; which gives a value of R_{DA}=1.89 and R_{DX}=1.42. 94

Table 7.2: Results of the FFBS data simulation. Data generated using the theoretically-generated Faraday filter functions as described in the text and analysis Faraday filter theoretical temperatures given in the table. Other generation and analysis parameters were the same as described in the text and in the heading for Table 7.1. 95

Table 7.3: Faraday filter *c* coefficients and γ values for data presented in Section 7.4.1. 99

Table 7.4: Summary of data and comparison to Slanger et al (2005)..... 105

Table 7.5: Observed nightly mean of $R_{D(A)}$, $R_{D(X)}$, R_{D1}, and R_{D2} results The number in parentheses next to each average is the number of points used to determine the average. The number in parentheses next to the duration is the total number of data points. 106

Table A.1: Matrix elements for the $^2P_{3/2}$ state of Na. 134

Table A.2: D_1 allowed transition F_1 value and polarization. 136

Table A.3: D_2 allowed transitions F_1 and F_2 values and polarization..... 138

Table C.1: Fitting parameters and errors for the data from 26 September 2009…….157

CHAPTER 1

INTRODUCTION

The atmosphere is a very dynamic system, and much is still being learned about it. This is perhaps most clearly understood by noting that the mesosphere and lower thermosphere (MLT) region was termed the "ignorosphere" until the late 1980's because so little was known about the dynamics (and chemistry) about this region. It is also true for other regions of the atmosphere as political debates rage on the potential effects of climate change and how to mitigate them, and vast scientific research efforts are expended in this area.

In the past two decades, progress has been made in our understanding of the MLT, including the sodium resonance lidar work at Colorado State University, so that the ignorosphere designation has been dropped. However, there are still many outstanding questions on short-period waves, chemistry and composition of minor species in the MLT that await investigation and research. It will be noted that there is still no direct way to regularly measure concentrations of atomic oxygen (a major energy transporter) in this region, and that aspects of the sodium chemistry cycle are still poorly understood. It is my hope that the results presented in this dissertation will help to increase knowledge about this very important aspect of the MLT.

To that end, this dissertation will describe the design, use, and results from a new instrument: the Faraday Filter-Based Spectrometer (FFBS), which measures the ratio of the intensity of sodium nightglow D_2 and D_1 emissions from the MLT as well as other properties for determining the atomic oxygen to diatomic oxygen concentration ratio. I will show that the results produced by this instrument can provide valuable, short (1 hour or less) time scale information about the variation in the atomic oxygen concentration—results which are necessary for a complete understanding of the chemistry of the MLT region. I will also show how making these nightglow measurements with the co-located CSU sodium temperature and wind lidar can provide new information about the tidal variability of both the sodium nightglow and the atomic oxygen concentration.

The remainder of this chapter will give an introduction to the atmosphere, including how the layers interact, with a special focus on the chemistry. Next, I will introduce a history of some of the techniques used for studying the MLT, including lidar and passive optical instruments. The chapter will conclude with an outline of the remaining chapters of the dissertation.

1.1 Atmospheric structure

The atmosphere can be thought of as four layers, which are defined by the sign of the slope of the annual mean neutral temperature with altitude. The thin regions between layers, at either a local minimum or maximum in temperature, is known as the "–pause" of the layer below it. This general structure is illustrated in figure 1.1. The layers are the troposphere, stratosphere, mesosphere, and thermosphere. Each of the major layers is described below.

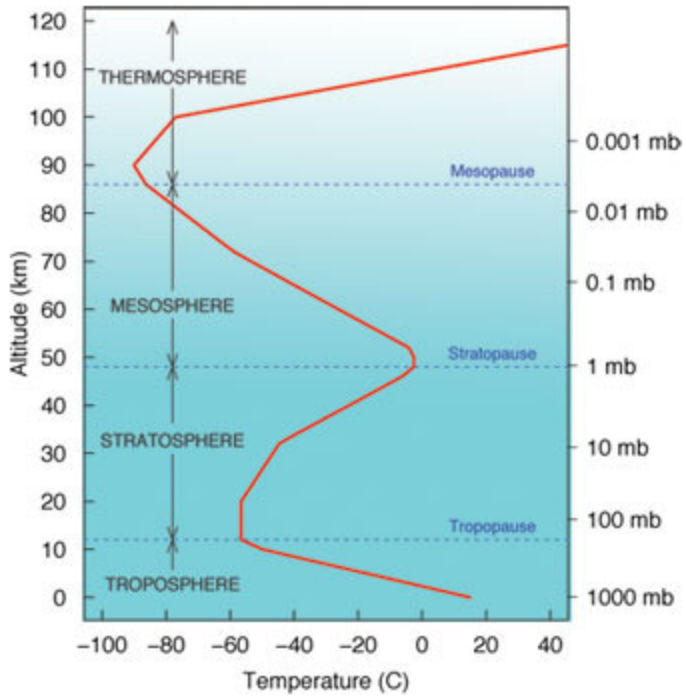

Figure 1.1: Schematic of atmospheric thermal variation as a function of both altitude and pressure, with layer designations. From www.learner.org.

The troposphere ranges from 0 to around 10 km (it can range as high as 16 km in the tropics), and the temperature decreases as altitude increases. This is due to the warming of the earth's surface, which re-radiates energy to warm the air near the surface. The troposphere contains most of the earth's weather and the majority of the total density of the atmosphere, which exponentially decreases (along with pressure) with altitude throughout the atmosphere. The major species are N_2 and O_2, which account for 78% and 21%, respectively, of the total density in the troposphere (Smith 2004). The tropopause, near 10 km in the polar region and up to 16 km in the tropics, is at the altitude of temperature minimum between the troposphere and stratosphere.

In the stratosphere, the temperature increases due to the presence of the ozone layer in the lower half of the stratosphere. This layer absorbs ultraviolet radiation and

reradiates in the infrared, which warms the surrounding air. The stratopause represents the maximum in temperature between the stratosphere and mesosphere.

The mesosphere, ranging from 50 km to 90 km, is somewhat of an enigma. Since there is no heating source, such as water vapor in the troposphere and ozone in the stratosphere, the temperature should decrease with altitude. In fact, the mesopause region, around 90 km, is the coldest region of the earth's atmosphere. However, the mesopause temperature at the solstices is not in radiative equilibrium—the summer solstice is colder than the winter solstice in the region near the mesopause. The explanation of this anomalous temperature structure lies with gravity wave filtering and breaking (Lindzen 1981). Gravity waves are the low-frequency equivalent of acoustic waves and have gravitational and buoyancy restoring forces (Hines 1960). Gravity waves are selectively filtered by lower atmospheric winds, and when they reach the mesosphere they increase in amplitude and break, depositing their momentum as a body force to reverse the mean zonal wind. This serves to create a meridional flow from the summer pole to the winter pole, cooling the summer (and warming the winter) mesopause (Holton 1983). This process is illustrated in Figures 1.2 and 1.3. The composition of the lower mesosphere is dominated by molecules. However, the lower mesosphere is also a transition region as some molecules that are dominant in the troposphere begin to break down due to photolysis. This means that daytime concentrations of O, H, and Cl atoms are high, but negligible during the night (Smith 2004).

The thermosphere ranges from 90 to 500 km in altitude and contains the warmest temperatures in the earth's atmosphere. The temperature increases with increasing altitude due to the rapid ionization from photolysis and auroral activity and photo

dissociation of the little remaining air density at these altitudes, meaning a decrease in the percentage of molecules and an increased percentage of atoms. Reactions are slow because of the low density and large mean free paths (Smith 2004). This ionization also leads to the co-located ionosphere, which ranges from 60 km in altitude to the upper reaches of the thermosphere and magnetosphere. This region is extensively studied with various radar techniques and makes possible radio communications on earth.

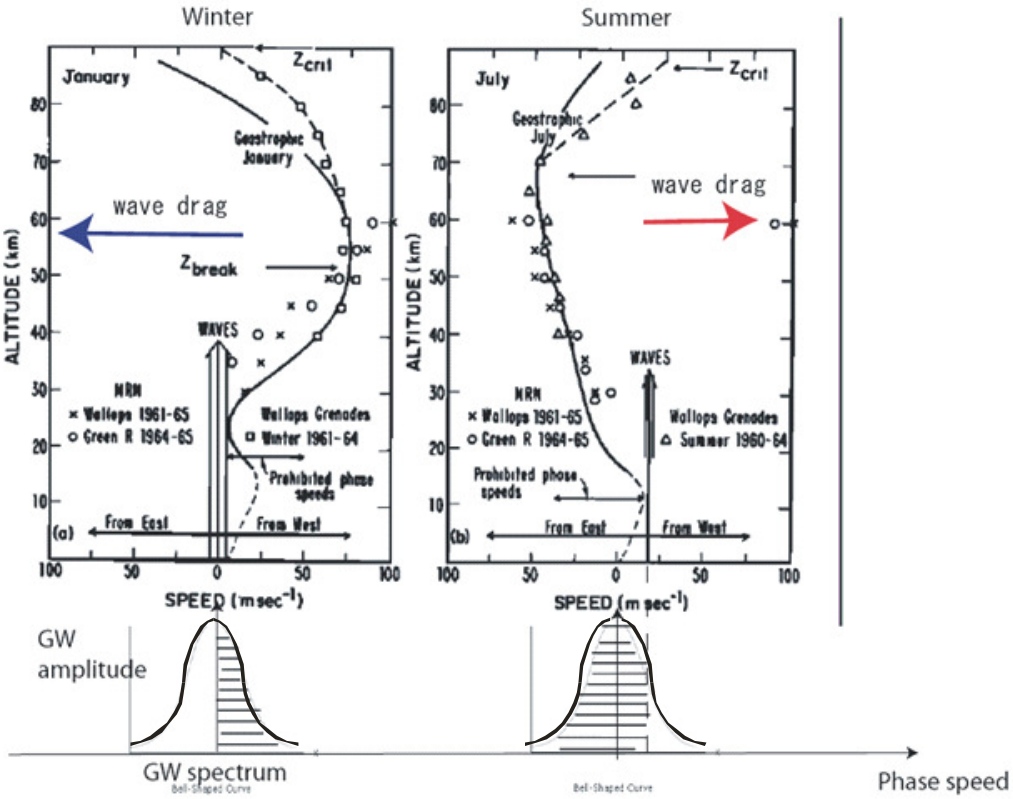

Figure 1.2: Plot of mean zonal wind in winter (left) and summer (right). The bottom of the plot shows the Gaussian distribution of gravity wave phase speeds; the shaded portion indicates the phase speeds that are filtered by the mean wind and therefore cannot transmission to the MLT. Those waves that do survive to the MLT break, and impart wave drag in the direction indicated. (from Lindzen 1981)

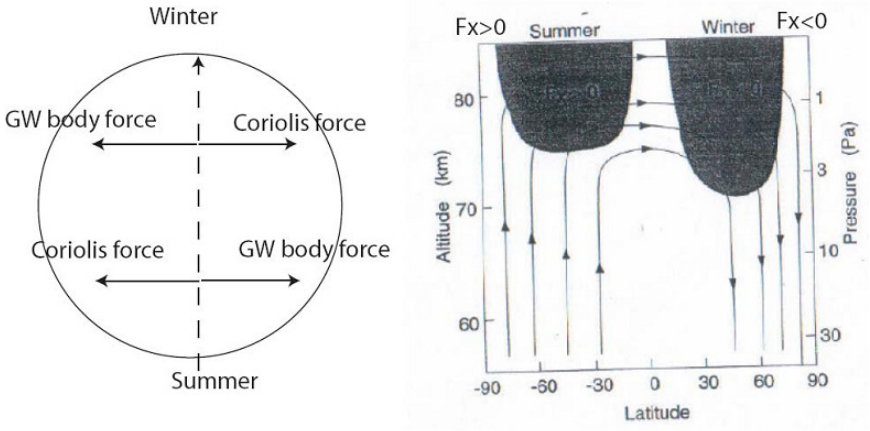

Figure 1.3 MLT region forcing. Left: the Coriolis force balances the gravity wave drag, or body force, to create a meridional flow from the summer to the winter pole in the MLT region (from Yue, 2009). Right: the meridional flow which leads to air rising in the summer, and falling in the winter, cooling the summer and warming the winter mesopause (from Holton and Alexander, 2000).

1.2 Mesosphere-lower thermosphere

The region near the mesopause, known as the mesosphere- lower thermosphere (MLT), has already been mentioned as both a poorly measured and enigmatic region of the atmosphere. This region will be the focus of this dissertation.

Early studies of the MLT region dynamics typically utilized rocket-based falling sphere measurements. While these experiments provide valuable information, they have low vertical spatial resolution and are extremely expensive and infrequent. In recent years, both ground-based and satellite-based techniques have been developed. Ground-based radar has become an effective probe for winds in the MLT. The medium frequency (MF) and very high frequency (VHF) radar depend on Bragg scattering from charged turbulence eddies and can be used to measure winds between 65 and 90 km (Sahr 2006). Meteor radar depends on scattering from meteors that can cover 65 to 100 km and can also measure winds (Palo 2007). In addition, meteor radar can infer atmospheric temperature by measuring meteor decay time.

Ground-based lidars typically use the metal layers (Na, Fe, K, and others) as fluorescence sources to measure metal density, temperature, and winds. These naturally occurring metal layers arise from the ablation of meteors which are constantly striking the earth's atmosphere. As the meteors enter the atmosphere, friction warms them and their metal constituents are vaporized. Below the MLT, however, the increased air density causes these neutral metals to quickly form molecular compounds with other atmospheric species. It is a fortuitous coincidence that the MLT region is still of a low enough density to avoid much formation of metal compounds while also being at an altitude which allows for many of the meteors to be ablated. This facilitates experiments which have increased our knowledge of the MLT region. More detail about the CSU Na lidar experiment will be given in chapter 6.

The MLT region can also be studied with other techniques besides those mentioned above. Fabry-Perot interferometers can be used to study the nightglow and determine MLT temperatures through investigating thermal Doppler broadening of some nightglow features (Sargoytchev, et al. 2004). Winds have been measured passively by the Magneto-Optic Doppler Analyzer (MODA), which utilized filtering of sodium nightglow emissions (Williams and Tomczyk 1996). Radar systems can also be utilized to study the MLT region; for example, they can study the charged ice particles in Polar Mesospheric Summer Echoes (She, et al 2006).

Another coincidence is the presence of nightglow. Nightglow is the non thermal radiation from the earth's atmosphere at night and will be detailed in chapter 2. Many of

the nightglow emissions, from species like Na, O_2, OI, and OH coincidently originate from the MLT region. These nightglow emissions, as mentioned above, have facilitated the study of the chemistry and dynamics of the MLT region.

1.3 Outline of dissertation

Chapter 2 of this dissertation will contain a more detailed description of MLT chemistry, with a focus on the chemistry of the oxygen species and sodium atoms. Previous attempts and other suggested techniques to measure atomic oxygen concentrations will be briefly presented.

Chapter 3 will give an introduction to nightglow, with a special focus on the sodium nightglow. This will include a discussion of the history of sodium nightglow measurements and theory, and the current state of understanding of the chemical production pathway of sodium nightglow. Various parameters used in this study will also be derived based on the steady-state chemistry of the modified Chapman mechanism.

I will present a complete description of the Faraday filter physics in chapter 4, which will facilitate a description of the FFBS instrument in chapter 5. Chapter 5 will include a discussion of the optical design, measurement technique, and the data analysis scheme.

Chapter 6 will contain a brief overview of the co-located CSU sodium Doppler lidar. This chapter will end with a description of how the lidar temperature and wind data can be used to further study the oxygen and sodium chemistry.

Spectrometer results will be presented in chapter 7. Specifically, the D_2/D_1 intensity ratio will be shown to be comparable to previous results from Slanger et al

(2005). The first known results of other parameters specific to the modified Chapman mechanism will also be presented.

The dissertation will conclude with general conclusions and suggestions for future work in chapter 8.

CHAPTER 2

MESOPAUSE REGION CHEMISTRY

The mesopause region, while having a very low density when compared to other atmospheric regions, has a large number of chemical constituents which are integrally tied with the dynamics of the region. For example, not only do a number of chemical reactions depend on ozone, it also absorbs UV solar radiation and is a major source of atmospheric heating and thus drives atmospheric tides. In order to fully understand the atmosphere, the chemistry must be understood as well. Atmospheric models must also include relevant chemical interactions in order to produce accurate results.

In this chapter, I will give a general review of mesopause region chemistry. I will focus on the role of atomic and diatomic oxygen and atomic sodium, since these will play an important role in the later parts of this dissertation. I will also describe various measurement techniques for atomic oxygen concentration and their advantages and disadvantages.

2.1 General Composition and Chemistry

The composition of the mesopause has been summarized by Smith (2004), and this section will summarize some of the important topics from this paper. The two most abundant molecules are O_2 and N_2—N_2 is little changed from its lower atmosphere

abundance of 78%, and O_2 is around 10% less in the thermosphere than it is in the troposphere. O_2 is specifically important because of the role it plays in absorbing solar radiation and as the source of more reactive oxygen species.

Reactions in the mesopause region consist of three types: photolysis, 2-body, and 3-body reactions. Photolysis is the splitting of a molecule into constituents upon absorption of a solar photon. Two-body chemical reactions involve two reactants and two products, while 3-body reactions contain two reactants, one product, and a third initial molecule needed to conserve momentum which is otherwise unchanged. The reaction rates involved in chemical reactions in the atmosphere are typically temperature dependent, connecting the chemistry and the dynamics.

The mesopause region composition is a combination of some aspects of the stratosphere and lower mesosphere below and the thermosphere above. In common with the lower mesosphere is that the atoms O_2 and N_2 are still well mixed—a fact important in the Faraday filter-based spectrometer measurements to be described later. Atomic species, such as O and H, at the mesopause are long-lived since low density reduces the chance of a reaction. The model results in figures 2.1 and 2.2 show the concentration profiles of several chemical constituents in daytime and nighttime, as well as various times of the year. These are expressed in terms of mixing ratio, which is defined as the number density of the constituent divided by the total number density. These figures demonstrate the diurnal and annual time scales upon which the composition varies; it also varies at longer time scales, such as the solar cycle.

The mean circulation of the atmosphere serves to mix constituents between atmospheric levels. For example, falling air masses in the winter carry constituents from

the thermosphere into the mesopause region, increasing concentrations of atomic species such as O. Tidal temperature variations can also affect reaction rates. Gravity wave-induced perturbations can also transport O downwards into the mesopause region.

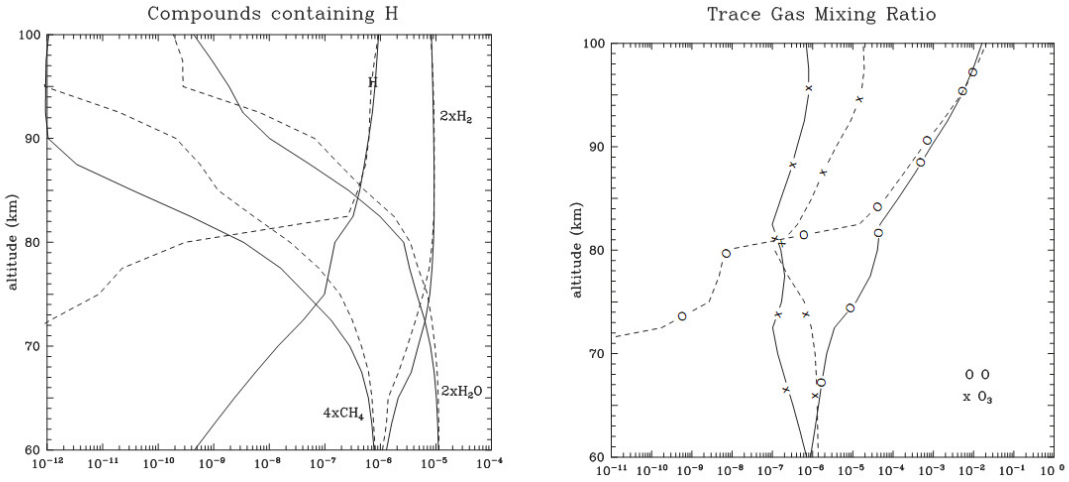

Figure 2.1: Mixing ratio profiles of hydrogen-containing compounds (left) and oxygen containing compounds (right) as calculated by the ROSE model. Hydrogen profiles are scaled by the number of hydrogen atoms they contain. Solid lines are for daytime, dashed lines for nighttime. From Smith (2004).

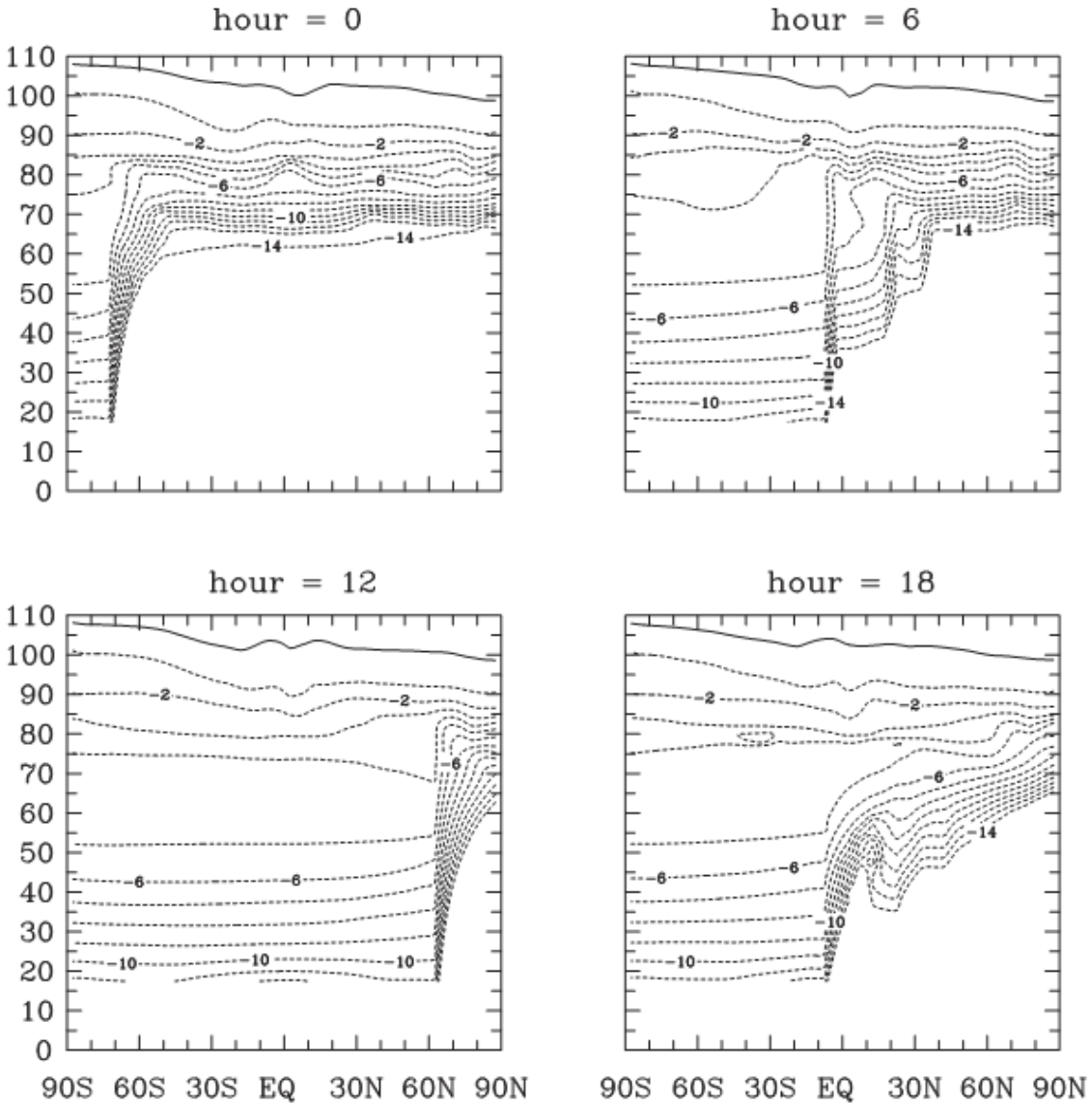

Figure 2.2: Logarithm of atomic oxygen mixing ratio as a function of latitude and height at four local times as calculated by the ROSE model. From Smith (2004).

2.2 Oxygen measurements: conclusions and limitations

While Smith mentions that it is difficult to measure O concentrations, some work has been done, and important conclusions have been drawn. Offermann *et al*. (1981) provide measurements with rocket based mass spectrometers, which measured a variety of constituents, including O. Yee *et al.* (1997) provide airglow measurements to determine the abundance of $O(^1S)$ (an excited state of oxygen). A two-photon lidar was

proposed by McIlrath *et al* (1979) which would have allowed for satellite-based lidar measurements of the vertical distribution of the oxygen concentration, but this lidar was never constructed. Satellite measurements have also been performed by several instruments, such as the WINDII instrument on the UARS satellite and the SABER and TIDI instruments on the TIMED satellite. Ground-based measurments using photometry of OI at 557.7 nm, O_2 atmospheric (0-1) band near 866 nm, and OH Meinel (9-4) band at 755 and 772 nm nightglow emissions have been performed by Melo *et al*. (2001)

Each measurment scheme discussed above has its advantages and disadvantages. Rocket-based measurements give altitude distributions directly, but their measurments are infrequent and expensive. The lidar measurment would be an ideal ground-based or satellite-based option to directly measure the oxygen concentration with both good altitude and time (in the case of ground-based lidar) coverage, however as mentioned it was never constructed. Satellite measurements give altitude dependence based on limb scanning and have global coverage, but they have poor local time coverage. Ground-based measurements with photometers are indirect, require parameterization to get altitude distrubutions, and there is only one set of published results (to my knowledge), showing only seasonally averaged parameters with inherently large error. (Melo et al 2001)

2.3 Importance of atomic oxygen

The importance of atomic oxygen in the mesopause region cannot be overstated. Thomas presented results from the Solar Mesospheric Explorer spacecraft which showed evidence of a seasonal variation of O, based upon a seasonally varying production rate by photolysis of O_2 (Thomas 1990). Smith emphasizes the importance of this photolysis

process, stating that the photolysis occurring in the thermosphere leads to a downward transport through diffusion or mean advection of O, where in the higher-density mesopause region it is able to react and release heat. The largest contributors to so-called chemical heating are reactions involving the O produced through photolysis from above. Atomic oxygen is also involved in the production of many of the airglow emissions.

Because of the chemistry and physics discussed in this chapter, an understanding of the short-term variations (on the order of one hour) of atomic oxygen at a particular location would be very useful for understanding the interactions between chemistry and dynamics in the mesopause region. While there have been attempts at making these types of measurements, more work is needed to understand the short term variations in atomic oxygen concentration. Therefore, a short-term measurement, even if indirect, would be quite useful. The Faraday Filter Based Spectrometer will provide such a measurement, and will allow for short-time scale determination of the variation of atomic oxygen concentration.

2.4 Chemistry of Mesospheric sodium

The sodium layer develops from the ablation of meteoroids in the earth's atmosphere. This process also leads to layers of other meteoric metals, such as Fe, Ca, and K, all near the height of the mesopause. Many of these species have been used for lidar experiments to determine MLT dynamics (see, for example, Gerding, et al. (2000), Hoffner and Lautenbach (2009), and Friedman, et al. (2003)). They also have unique chemical interactions that allow them to be used for studies of the chemistry of the mesopause region.

Plane (2003) reviews the chemistry of these metal species in the atmosphere. He states that meteoroid input from two sources: dust trails produced by sublimating comets orbiting the sun, and fragments from the asteroid belt and dust from decayed comet tails. The second source gives a reasonably constant source of metals. The daily mass flux is an uncertain quantity; various estimates place it in the range of 50 to 250 tons per day, with sizes on the order of μm and entry speeds on the order of tens of km s^{-1}. The physical process of ablation can be understood for particles less than 250 μm. From this theory, it is understood that with these high entry velocities, meteoroids will rapidly heat through collisions with air molecules, and their constituents will vaporize; Na vaporizes at 1000 K, and all meteoroids larger than 10 μm will reach this temperature.

The density of sodium has a peak between 88 and 92 km altitude, with a FWHM thickness of about 10 km. This position and thickness (as well as the shape of the layer) is dictated by chemistry—Plane states that above the Na layer, Na atoms become ionized by interactions with NO^+ and O_2^+. Below the layer, Na is converted into $NaHCO_3$, which is a stable reservoir. Reactions with O and H are needed to convert the $NaHCO_3$ back to atomic Na, so the Na layer only exists in regions with relatively high concentrations of O and H. Figure 2.3 schematically shows this chemical cycle of Na. Several of the reactions in the figure have important implications for Na nightglow production (to be detailed in Chapter 3). While the nightglow is driven by the production of NaO via a reaction of Na and O_3, O is necessary to complete the process and produce a photon and a ground state Na atom. If the concentration of the atomic oxygen is too low, however, the NaO will react with H_2O and O_2 to form the more stable reservoir species of NaO_3 and NaOH (Plane, Personal communication 2010).

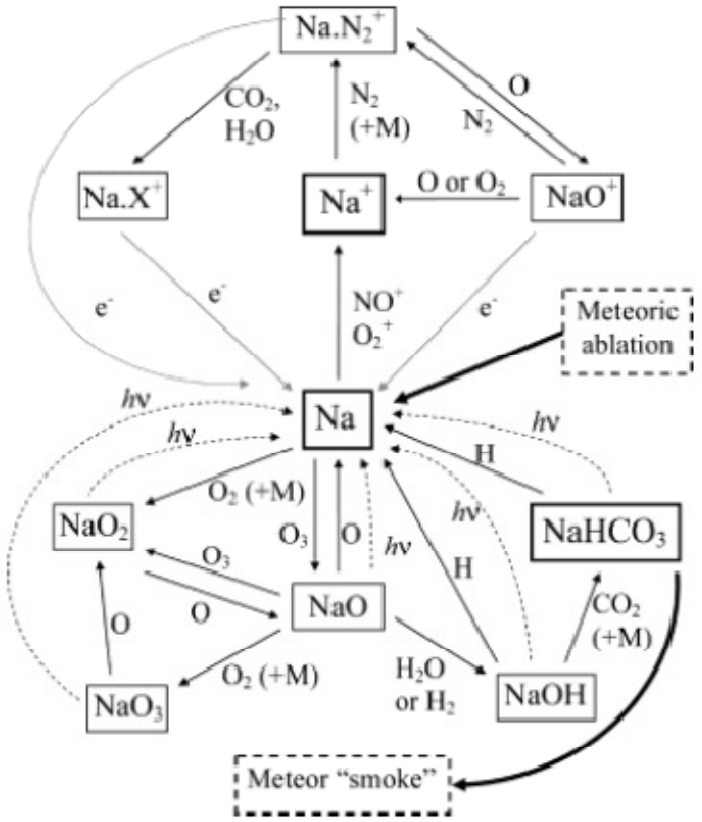

Figure 2.3: Schematic showing significant chemical cycles of sodium in the mesopause region.Input of sodium comes from meteoric ablation, and sodium is removed from the system by the formation of meteor smoke, which is metallic compounds polymerizing together with silicon oxides, which condense and coagulate into particles which are too large to remain suspended above 80 km. From Plane (2003).

CHAPTER 3

NIGHTGLOW

Airglow is non-thermal radiation emitted by earth's atmosphere, excluding aurora and lightning (Chamberlain 1961). It is created through chemical processes, and can be produced through atomic or molecular transitions in many species, including Na, O, O_2 and OH. Nightglow is a subset of airglow, which can also include resonant scattering of sunlight as an excitation mechanism during the daytime (dayglow). In this chapter, I will begin with a brief introduction of nightglow. This will be followed by a detailed discussion of the production, properties, and relevant previous studies of the sodium nightglow.

3.1 History

Chamberlain (Chamberlain 1961) provides detailed discussion of the history of airglow observations, which will be summarized here. The existence of airglow was first established by Yntema in 1909, who gave it the name "earthlight" (Yntema 1909). However, he noted that observers had noticed nights of increased sky brightness for several centuries prior. The terms "airglow" and "nightglow" were not introduced until the 1950's. Rayleigh, using a variety of methods, investigated the nightglow spectrum and the intensity variations with time as well as relative intensities of various parts of the

spectrum (see, for example, Rayleigh (1924)). The unit used in photometry of the nightglow is named "rayleigh" in his honor; it is defined (for the purposes of this dissertation) as 1R= $(1/4\pi)10^{10}$ photons s^{-1} m^{-2} sr^{-1} (Baker and Romick 1976).

The sodium nightglow at the D line (589 nm) was first reported by Slipher (1929). Further work was done by Roach and Petit (1951) and Wiens and Weill (1973) on the annual and solar-cycle variations of the Na nightglow, and collocated lidar and nightglow measurements were made by Clemesha et al (1993). Later, the Na nightglow was used to determine mesopause region winds (Williams and Tomczyk 1996).

3.2 Sodium Nightglow Production I: The Chapman Mechanism

Several excitation mechanisms for nightglow have been suggested. The first was a proposal by Chapman (1939). While other production schemes were suggested, and it was initially difficult to validate Chapman's proposal (Plane 2003), laboratory work and nightglow measurements have validated most aspects of the Chapman mechanism (Sipler and Biondi 1978; Joo, et al. 1999; Shi, et al. 1993).

The sodium nightglow arises from the ozone chemistry-induced excitation of mesospheric sodium atoms (Chamberlain 1961). This process is known as the Chapman Mechanism, and excited states of sodium are produced that lead to nightglow at two wavelengths: D_2 (589.158 nm) and D_1 (589.756 nm):

$$\text{(R1)}\quad \mathrm{Na} + \mathrm{O_3} \rightarrow \mathrm{NaO} + \mathrm{O_2}$$

$$\text{(R2)}\quad \mathrm{NaO} + \mathrm{O} \rightarrow \begin{cases} \mathrm{Na}(^2\mathrm{S}_{1/2}) + \mathrm{O_2} \\ \mathrm{Na}^*(^2\mathrm{P}_{1/2}) + \mathrm{O_2} \\ \mathrm{Na}^*(^2\mathrm{P}_{3/2}) + \mathrm{O_2} \end{cases}$$

$$\text{(R3)}\quad \mathrm{Na}^*(^2\mathrm{P}_{3/2}) \rightarrow \mathrm{Na}(^2\mathrm{S}_{1/2}) + I(D_2 = 589.158\,\text{nm vacuum})$$

$$\text{or}\quad \mathrm{Na}^*(^2\mathrm{P}_{1/2}) \rightarrow \mathrm{Na}(^2\mathrm{S}_{1/2}) + I(D_1 = 589.756\,\text{nm vacuum})$$

There are several observations related to these reactions. First, there is no *a priori* value for the intensity ratio of D_2/D_1 (Slanger and Plane 2005). If the excited states are populated according to the statistical weights of their spin-orbit coupling, then $R_D = I(D_2)/I(D_1)$ will be 2. However, the chemical reactions involved in the Chapman Mechanism may or may not populate the excited states in this way. In addition, the mean time between collisions at the mesopause (~0.04 ms) is greater than the ^2P lifetime (~16 ns), so the Na atoms will not equilibrate with the atmosphere before a photon is emitted. Also, the lineshape of the emission at both lines will be broader than the sodium spectrum without chemical reactions at the ambient atmospheric temperature —this is due to the dissociation kinetic energy, which is imparted to the sodium atom in reaction R2 from the break-up of the NaO molecule. This will be detailed in section 3.6.

3.3 Measurements of Sipler and Biondi

The measurements of Sipler and Biondi (1978) put to rest debate over the value of D_2/D_1, at least temporarily. Their measurements were made with a 150 mm aperture, pressure scanned Fabry-Perot interferometer, which gave values for the intensity ratio as well as the linewidth of the sodium transition. These measurements gave a value for R_D of 1.98±0.1 as shown in Figure 3.1; until recently R_D was assumed to have a "fixed" value of 2. The measurements also showed a semiannual variation in the nightglow linewidth with higher values in May and November, and the total range of linewidth values from 282 cm^{-1} and 403 cm^{-1}. This led them to accept the Chapman mechanism as the probable source of the Na nightglow.

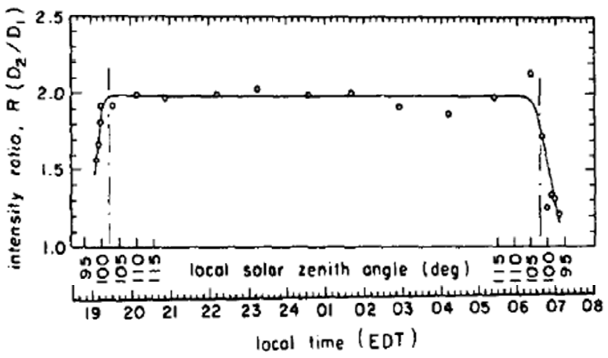

Figure 3.1: Nightglow results from Sipler and Biondi, taken with a scanning Fabry-Perot Interferometer. The average value of their nightglow measurements is 1.98±0.1. The times to the left and right of the dashed lines (local solar zenith angle > 100°) is during twilight. Twilight airglow becomes dominated by resonant florescence by sodium of solar photons; since the D_2 is 90% of the D_1 emission from the sun, coupled with the stronger D_2 absorption of sunlight at D_2 frequencies as the light traverses the sodium layer to reach the observation point, the observed R_D quickly decreases (Sipler and Biondi 1977).

3.4 Sodium nightglow production II: Measurements of Slanger *et al.* and the modified Chapman Mechanism

The results of Sipler and Biondi were accepted until 2005, when the studies of Slanger *et al*. (2005) cast doubt on the static value of R_D. These results, using échelle spectrographs on the Keck I and II telescopes, showed a variation of R_D during the year, with peaks at the equinoxes and minima at the solstices, and values between 1.2 and 1.8, as shown in figure 3.2. Slanger et al. made several attempts to explain this somewhat surprising result by using laboratory investigations of the Chapman Mechanism. Their experiments utilized a fast flow tube reactor to simulate the conditions in the mesopause region. They tested the dependence of R_D on temperature, the relative concentration of O and O_3, and concentrations of the possible quenching molecules N_2 and O_2. The only one of these experiments which showed any variation of R_D was the presence of the quenching molecule O_2. They also present arguments for the elimination of self

absorption in the Na layer and a completely different nightglow production mechanism (such as recombination of Na^+ cluster ions and exothermic heterogeneous reactions on the surface of metoric smoke particles) as possible causes of the variation of R_D. The only possible solution, they believe, is a modification of the Chapman mechanism.

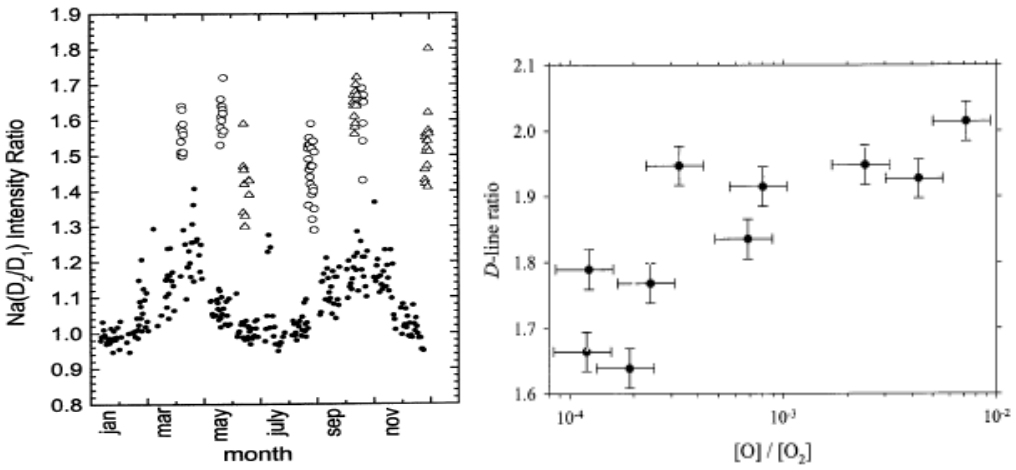

Figure 3.2: Results from Slanger *et al*, showing annual variation in R_D (open points) and nightglow intensity (solid points) (left pane) and R_D vs. $[O]/[O_2]$ (right pane). For details, see Slanger et al. (2005).

Reaction (R2) can proceed as above, with the NaO molecule created by (R1) in the $(A^2\Sigma^+)$ excited state:

$$\text{(R1)}\quad Na + O_3 \xrightarrow{k_1} NaO + O_2$$

$$\text{(R2)}\quad NaO^*(A^2\Sigma^+) + O(^3P) \xrightarrow{k_2} \begin{cases} p_{sA}\,Na(^2S_{1/2}) + O_2 \\ p_{1A}\,Na^*(^2P_{1/2}) + O_2 \\ p_{2A}\,Na^*(^2P_{3/2}) + O_2 \end{cases}.$$

For simplicity, the $(A^2\Sigma^+)$ state will now be simply labeled (A), and all parameters involved in this reaction will be labeled with a subscript *A*. The reaction rates are labeled k_1 and k_2 respectively. Each state of sodium has a fractional probability of being produced: p_{sA} is the probability of producing the $Na(^2S_{1/2})$ ground state, p_{1A} is the probability of producing the $Na(^2P_{1/2})$ excited state (which produces D_1 emission), and

p_{2A} is the probability of producing the Na($^2P_{3/2}$) excited state (which produces D_2 emission). The NaO(A) can also be quenched to the ground state by a quenching molecule *M*:

$$\text{(R2a)}\quad NaO^*(A^2\Sigma^+) + M \xrightarrow{k_{2a}} NaO(X^2\Pi) + M$$

$$\text{(R2b)}\quad NaO(X^2\Pi) + O(^3P) \xrightarrow{k_{2b}} \begin{cases} p_{sX}\, Na(^2S_{1/2}) + O_2 \\ p_{1X}\, Na^*(^2P_{1/2}) + O_2 \\ p_{2X}\, Na^*(^2P_{3/2}) + O_2 \end{cases}$$

We will abbreviate NaO($X^2\Pi$) as NaO(X), and a subscript *X* will indicate parameters belonging to this reaction. The sodium state probabilities and the reaction rates are defined similarly to (R2). Note that by definition, $p_{sA}+p_{1A}+p_{2A} = p_{sX}+p_{1X}+p_{2X} = 1$.

It should be noted that the various probabilities *p* are not necessarily the same. Also, laboratory work has shown that (R1) produces NaO entirely in the (A) excited state (Wright 1993; Shi et al. 1993), and further discussion of the plausibility of this reaction sequence is discussed by Slanger et al. (2005). The most likely candidates for *M* are O_2 or N_2. Results of laboratory experiments presented in Slanger et al. (2005) show a dependence of R_D on [O]/[O_2], but no dependence on [O]/[N_2], as shown in figure 3.2. Further evidence for O_2 quenching comes from energy considerations; while (R2a) with N_2 is endothermic by 411 cm^{-1}, the reaction with O_2 is exothermic by 363cm^{-1}. These support the assumption that *M* is O_2.

Both (R2) and (R2b) are exothermic, however by differing amounts of released energy. This extra energy causes a preferential population of the $^2P_{3/2}$ state for (R2), and the $^2P_{1/2}$ state for (R2b). These two competing pathways (R2 and R2a-b) will occur at the same time; as the relative values of [O] and [O_2] vary, the observed values of R_D and their contributions from (R2) and (R2b) will vary as well.

As previously stated, reactions (R2) and (R2b) are both exothermic, but they release different amounts of energy because of the different NaO states at play in each reaction. This extra energy will impart a constant velocity to the sodium atoms released, which is manifested as a broadening of the spectrum emitted (Rogers and Biondi 1964). This broadening will be determined in section 3.6. Because more energy is released in (R2), the D-line spectrum produced will have broader lines than those produced from (R2b), while both are broader than the spectrum predicted from the thermalized hyperfine structure. It should be noted that reactions (R2) and (R2a) provide the extreme cases; as the ratio $[O]/[O_2]$ varies, the lineshape of the spectral lines will be a weighted average dependent on the proportion of the nightglow produced by each reaction.

3.5 Sodium nightglow production III: Chemical rate equations and implications for nightglow measurements

To study the nightglow and the related chemistry, we must solve the rate equations for the production of nightglow. First, reaction (R3) must be rewritten to accommodate the two reaction pathways:

$$\begin{aligned}
&(\mathrm{R3}(\mathrm{D_{2A}})) \quad \left(\mathrm{Na}(^2\mathrm{P_{3/2}})\right)_A \xrightarrow{k_3=1/\tau} \mathrm{Na}(^2\mathrm{S_{1/2}}) + \mathrm{I_{2A}} \\
&(\mathrm{R3}(\mathrm{D_{1A}})) \quad \left(\mathrm{Na}(^2\mathrm{P_{1/2}})\right)_A \xrightarrow{k_3=1/\tau} \mathrm{Na}(^2\mathrm{S_{1/2}}) + \mathrm{I_{1A}} \\
&(\mathrm{R3}(\mathrm{D_{2X}})) \quad \left(\mathrm{Na}(^2\mathrm{P_{3/2}})\right)_X \xrightarrow{k_3=1/\tau} \mathrm{Na}(^2\mathrm{S_{1/2}}) + \mathrm{I_{2X}} \\
&(\mathrm{R3}(\mathrm{D_{1X}})) \quad \left(\mathrm{Na}(^2\mathrm{P_{1/2}})\right)_X \xrightarrow{k_3=1/\tau} \mathrm{Na}(^2\mathrm{S_{1/2}}) + \mathrm{I_{1X}} \quad .
\end{aligned}$$

While this looks confusing, the various iterations of reaction (R3) are necessary to account for the different nightglow intensities that may be emitted by each reaction. The reaction rate for all of the (R3) is k_3, which is the inverse lifetime of the sodium excited states (~16 ns for both excited states). I represents the nightglow intensity for the

respective pathways. Assuming that (R1), (R2), (R2a), (R2b) and (R3) are elementary reactions (Bromberg 1984), and noting that the reactions are fast (Plane 2003), then the rate equations of the intermediate reactants NaO(A), NaO(X), Na($^2P_{3/2}$) and Na($^2P_{1/2}$) can be solved for in the steady-state. The rate equations for these species are:

$$\begin{aligned}
\frac{d[\mathrm{NaO(A)}]}{dt} &= 0 = k_1[\mathrm{Na}][\mathrm{O_3}] - k_2[\mathrm{NaO(A)}][\mathrm{O}] - k_{2b}[\mathrm{NaO(A)}][\mathrm{O_2}] \\
\frac{d[\mathrm{NaO(X)}]}{dt} &= 0 = k_{2a}[\mathrm{NaO(A)}][\mathrm{O_2}] - k_{2a}[\mathrm{NaO(X)}][\mathrm{O}] \\
\frac{d\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_A}{dt} &= 0 = k_2 p_{2A}[\mathrm{NaO(A)}][\mathrm{O}] - k_3\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_A \\
\frac{d\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_A}{dt} &= 0 = k_2 p_{1A}[\mathrm{NaO(A)}][\mathrm{O}] - k_3\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_A \\
\frac{d\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_X}{dt} &= 0 = k_2 p_{2X}[\mathrm{NaO(X)}][\mathrm{O_2}] - k_3\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_X \\
\frac{d\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_X}{dt} &= 0 = k_2 p_{1X}[\mathrm{NaO(X)}][\mathrm{O_2}] - k_3\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_X
\end{aligned} \quad . \tag{3.1}$$

Solving the rate equations in (3.1) yields:

$$\begin{aligned}
[\mathrm{NaO(A)}] &= k_{2a}Q \\
[\mathrm{NaO(X)}] &= \frac{k_{2a}Q}{k_{2b}[\mathrm{O}]} \\
\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_A &= \frac{p_{1A}k_2Q[\mathrm{O}]}{k_3} \\
\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_A &= \frac{p_{2A}k_2Q[\mathrm{O}]}{k_3} \\
\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_X &= \frac{p_{1X}k_{2a}Q[\mathrm{O_2}]}{k_3} \\
\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_X &= \frac{p_{2X}k_{2a}Q[\mathrm{O_2}]}{k_3}
\end{aligned} \quad . \tag{3.2}$$

where

$$Q = \frac{k_1[\mathrm{Na}][\mathrm{O_3}]}{(k_2[\mathrm{O}] + k_{2a}[\mathrm{O_2}])} .$$

The rate of sodium nightglow production is:

$$\begin{aligned}
\frac{dI_{1X}}{dt} &= k_3\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_X \\
\frac{dI_{2X}}{dt} &= k_3\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_X \\
\frac{dI_{1A}}{dt} &= k_3\left[\mathrm{Na}(^2\mathrm{P}_{1/2})\right]_A \\
\frac{dI_{2A}}{dt} &= k_3\left[\mathrm{Na}(^2\mathrm{P}_{3/2})\right]_A
\end{aligned} \tag{3.3}$$

By substituting (3.2) into (3.3), and assuming the nightglow is measured over some interval of time Δt much longer than the reaction times, the four total nightglow intensities can be calculated:

$$\begin{aligned}
I_{1A} &= Qp_{1A}k_2[\mathrm{O}] \\
I_{2A} &= Qp_{2A}k_2[\mathrm{O}] \\
I_{1X} &= Qp_{1X}k_{2a}[\mathrm{O}_2] \\
I_{2X} &= Qp_{2X}k_{2a}[\mathrm{O}_2].
\end{aligned} \tag{3.4}$$

From equation (3.4), several parameters relevant to the modified Chapman mechanism of sodium nightglow can be derived:

$$\begin{aligned}
R_{D1} &= \frac{I_{1A}}{I_{1X}} = \frac{p_{1A}k_2[\mathrm{O}]}{p_{1X}k_{2a}[\mathrm{O}_2]} \\
R_{D2} &= \frac{I_{2A}}{I_{2X}} = \frac{p_{2A}k_2[\mathrm{O}]}{p_{2X}k_{2a}[\mathrm{O}_2]} \\
R_{D(X)} &= \frac{I_{2X}}{I_{1X}} = \frac{p_{2X}}{p_{1X}} \\
R_{D(A)} &= \frac{I_{2A}}{I_{1A}} = \frac{p_{2A}}{p_{1A}} \\
R_D &= \frac{I_{2X}+I_{2A}}{I_{1X}+I_{1A}} = \frac{\left(p_{2X}k_{2a} + p_{2A}k_2\dfrac{[\mathrm{O}]}{[\mathrm{O}_2]}\right)}{\left(p_{1X}k_{2a} + p_{1A}k_2\dfrac{[\mathrm{O}]}{[\mathrm{O}_2]}\right)}.
\end{aligned} \tag{3.5}$$

Note that both R_{D2} and R_{D1} are directly proportional to [O]/[O_2], and R_{DA} and R_{DX} give the D_2/D_1 intensity ratio for the individual reactions R2 and R2b.

3.6 Sodium nightglow production IV: Reaction broadened nightglow spectra

Because of the long mean time between collisions (~0.04 ms) in the low-density mesopause region as compared to the lifetime (~16 ns) of the ^{2}P excited state of Na, the Na atoms produced by reactions (R2) or (R2b) do not have time to thermalize with the environment before a photon is emitted. Due to negligible momentum of the visible photon, the transition of the Na atom from ^{2}P to $^2S_{1/2}$ will not affect the velocity of the atom. Thus, the spectra will not be the thermal broadened hyperfine spectrum. Though the center-of-mass of the NaO+O reaction system may move randomly with respect to the earth at the background temperature, the Na atom will move away from the center of mass of the reaction system with much higher speed depending on the kinetic energy released in the reaction (Rogers and Biondi 1964).

In this calculation, I will assume that the center of mass of the NaO+O reaction system can be described by a Boltzmann distribution. This NaO + O system will be abbreviated NaO_2. The distribution of speeds of the reaction system along the observation line of sight (the "lab" frame) or z axis will be:

$$P_{NaO_2}(v) = \left(\frac{M_{NaO_2}}{2\pi k_B T}\right)^{1/2} \exp\left(-\frac{M_{NaO_2}}{2k_B T} V_{NaO_2,z}^{\ 2}\right). \tag{3.6}$$

where P_{NaO_2} is the velocity probability distribution; M_{NaO_2} = 55 amu, the mass of the NaO_2 system; k_B is the Boltzmann constant; T is the temperature, and $V_{NaO_2,z}$ is the velocity of the NaO_2 product in the line of sight direction z.

When the excited state Na is produced in (R2) and (R2b), it will also have kinetic energy E_D imparted to it from the dissociation of the NaO. Therefore, the Na will be moving with an additional speed equal to $V_{Na} = (2E_D / m_{Na})^{1/2}$ in the center of mass frame of the NaO+O. There will be no preferred direction for this additional motion; but v_{Na} will have a particular component along the line of sight direction z as measured from the ground. Let θ be the angle between lab and the emission direction of the Na atom as measured from the center of mass.

In order to get the lineshape of the Na fluorescence, the Boltzmann velocity distribution (equations 3.6) of the center of mass and the velocity of the Na atom relative to the line of sight must be multiplied together along with the natural lineshape:

$$g_{nat}(\nu) = \frac{\Delta\nu_h}{2\pi\left[(\nu-\nu_0)^2 + (\frac{\Delta\nu_h}{2})^2\right]} \tag{3.7}$$

where $\Delta\nu_h$ is the natural linewidth. The resonant frequency in the natural linewidth will be Doppler shifted in the lab coordinates, so that ν_0 becomes

$$\nu_0 + \nu_0 \frac{V_{NaO_2}}{c} + \nu_0 \frac{V_{Na}\cos\theta}{c}, \tag{3.8}$$

in the lab coordinates. By combining equations 3.6, 3.7, 3.8 and the modified ν_0, the total reaction-broadened linewidth can be written:

$$g_{react}(\nu) = \frac{\lambda}{2}\left(\frac{M_{NaO_2}}{2\pi k_B T}\right)^{1/2} \int_{-1}^{1}\int_{-\infty}^{\infty} \exp\left(-\frac{M_{NaO_2}}{2k_B T} V_{NaO_2,z}{}^2\right) \frac{\Delta\nu_h\, d(\cos\theta)\, dV_{NaO,z}}{2\pi\left[(\nu-\nu_0-\nu_0\frac{V_{NaO_2,z}}{c}-\nu_0\frac{V_{Na}\cos\theta}{c})^2 + (\frac{\Delta\nu_h}{2})^2\right]} \tag{3.9}$$

The integral over $d(\cos\theta)$ is included since the Na atom can be released with equal probability with any angle θ to the line of sight z. The integral over $dV_{NaO,\,z}$ is the Voigt

function. The integral of the Voigt function over $d(\cos\theta)$ can be done numerically. It can also be integrated in the other order: The integral over $d(\cos\theta)$ is straightforward, and the integral over dv_{NaO} can be done using Gauss-Hermite quadrature. Both yield the same results.

3.6.1 Specific values for the modified Chapman Mechanism

For the case of the modified Chapman mechanism, the dissociation kinetic energy can be estimated using conservation of energy and momentum equations. The bond energy of NaO(X) has been calculated to be 21219 ± 688 cm^{-1} (Plane, Personal Communication 2010), the bond energy of O_2 is 41184 cm,$^{-1}$. Therefore, reaction (R2b) is exothermic with a dissociation kinetic energy of 3007 cm^{-1}. Since the NaO(A) electronic state is 1919 cm^{-1} above NaO(X), the dissociation kinetic energy in (R2) is 4929 cm^{-1}. The kinetic energy may be imparted to the products as translational energy of the Na or O_2, or rotational or vibrational energy of the O_2. One vibrational quantum of O_2 corresponds to 1556 cm^{-1} of energy; the released kinetic energy for (R2) can excite up to three vibrational quanta of O_2, and one for (R2b). Since there is one diatomic molecule reactant (NaO) and one product (O_2) with similar rotational constants in reactions (R2) and (R2b), and total angular momentum must be conserved, I assume that additional rotational excitation of O_2 in these reactions is not significant. That is, all the available reaction energy is partitioned into translational and/or vibrational excitation of the products. Table 3.1 lists the energies and corresponding velocities for the Na atom for all the possible cases of O_2 vibrational excitation, along with the temperature of the Boltzmann distribution with an equivalent average speed for reference.

Table 3.1: Na atom energies and velocities for different possible O_2 vibrational excitations and corresponding Boltzmann temperatures.

O_2 vibrational quanta in (R2)	O_2 vibrational quanta in (R2b)	Na atom energy (cm^{-1})	Na atom velocity (m/s)	Corresponding Boltzmann temperature (K)
0		2867	1728	3245
	0	1749	1350	1980
1		1962	1430	2220
	1	844	938	955
2		1057	1050	1196
3		151	397	171

3.6.2 Discussion

Although there are several possible choices of Na velocities for each reaction presented in table 3.1, I will assume here that there is one vibrational quanta of O_2 in both (R2) and (R2b); the reason behind this assumption is based on the fact that the probability for one vibrational excitation should be much higher in the basic reaction with one collision considered here. Assuming the ambient temperature is 200K, these values lead to the emission spectra shown in figure 3.3. The difference in FWHM linewidth between the spectra is small, only 1.7 GHz (0.002 nm), requiring a spectrometer with very high resolution to detect the difference. It should be noted that reactions (R2) and (R2b) provide the extreme cases; as the ratio $[O]/[O_2]$ varies, the net measured lineshape of the spectral lines will be a weighted average.

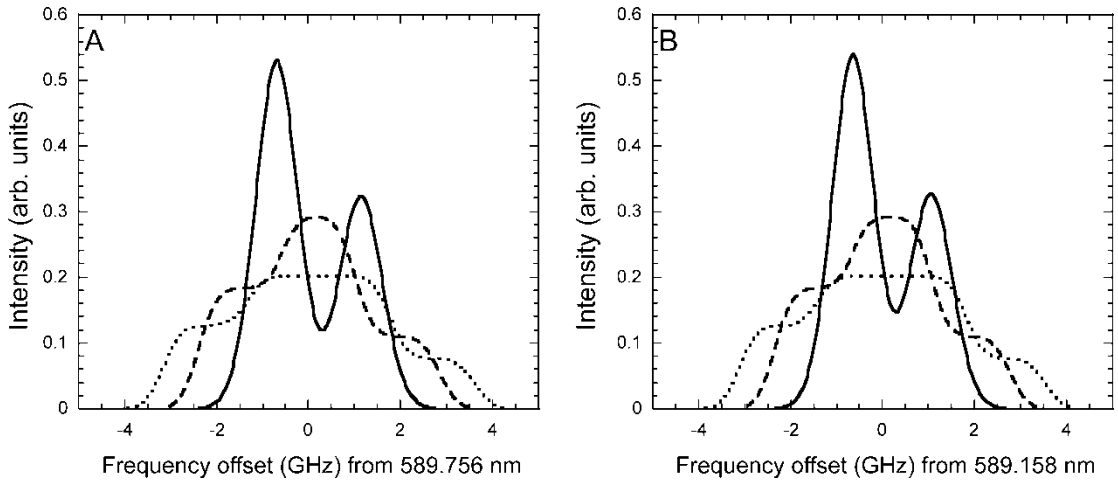

Figure 3.3: Calculated normalized nightglow spectra for (A) D_1 and (B) D_2. The dotted curve in each figure represents the spectra produced by (R2); the dashed curve is that from reaction (R2b). This curve is for the case of one vibrational quantum of O_2 for both (R2) and (R2b). Also plotted in solid for comparison are the normalized spectra as predicted by the quantum mechanical hyperfine structure. The temperature for each curve is 200 K.

As mentioned, there could be more than one vibrational excitation and/or an unknown amount of rotational excitation of the O_2 product of (R2) and (R2b), which was ignored in the above calculation. If this is not the case, and there is vibrational and/or rotational excitation of O_2 in reactions (R2) and (R2b), then the lineshapes could be quite different from those depicted in figure 3.3. This may indeed be the case; Fabry-Perot measurements of the nightglow lineshape by Hernandez (1971) yielded a value of 70±10 meV for a single observation. Further Fabry-Perot results from Sipler and Biondi (1978) reported measurements between 35 and 50 meV, and suggested a semi-annual oscillation of the linewidth with high values in May and November. Further laboratory experiments would need to be performed to validate the proper lineshape of the two reactions. The effects of this uncertainty on the FFBS nightglow measurements, however, are limited; Chapter 7 will detail why this is the case.

In equation (3.6), it was assumed that the center of mass NaO_2 could be described by a Boltzmann distribution at the ambient atmospheric temperature. Since the NaO_2 will only exist for a few femtoseconds (Plane, Personal Communication 2010), it is unlikely that it is at the ambient temperature. The choice of reactants for this initial Boltzmann distribution is not clear from the literature (see Sipler and Biondi (1978) and Hernandez (1975) for a discussion of the sodium lineshape, and Rogers and Biondi (1964) and Hernandez (1971) for a derivation of the similar lineshape for electron-ion recombination). In reality, the initial assumption of a moving center of mass is not necessary. Figure 3.4 shows sample spectra for two cases: the case of a 200 K and 50 K Boltzmann distribution of NaO_2. The figure indicates that the difference between these two cases is quite small and can be ignored for the analysis of nightglow measurements. Assuming that the proper atmospheric temperature Boltzmann distribution belongs to a different reactant (i.e., NaO or Na) would yield even smaller changes than those shown in figure 3.4.

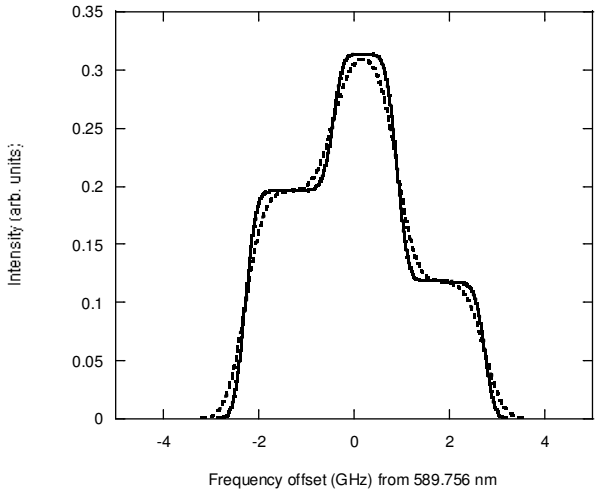

Figure 3.4: Difference between D_1 spectra produced by assuming a NaO_2 Boltzmann distribution of 50 K (solid) and 200 K (dashed). The spectra shown are for the NaO(X) or (R2b) case.

3.7 Necessity of the Faraday Filter-Based Spectrometer

It is obvious that a detection of the difference between Na nightglow produced by the different pathways of the modified Chapman mechanism requires a very sensitive spectrometer, since the difference in FWHM linewidth of the spectra produced by (R2) and (R2b) is small, only 1.7 GHz (0.002 nm). This cannot be achieved with any of the other instruments previously used to study the nightglow (Clemesha, et al. 1993; Hecht, et al. 1986; Slanger, et al. 2005; Sipler and Biondi 1978). Thus, it is clear a different approach is needed. The approach I will detail in the following chapters is based on narrowband optical Faraday filters. As will be described in Chapter 4, Faraday filters have a transmission bandwidth on the order of the nightglow linewidth, so that the filter's fractional transmission of the spectrum will be sensitive to the small changes in the proportion of nightglow produced by R_{D2} and R_{D1} which can allow determination of [O]/[O_2]. Chapters 4 and 5 will detail the equipment and the technique needed to determine all of the quantities in equation (3.5).

CHAPTER 4

FARADAY FILTER PHYSICS

Extremely narrow band optical filters are required in a variety of situations to extract useful signals in the presence of a broadband background. For detection of signals at an atomic resonance, a big advantage is attained by using a Faraday filter. Advantages of the atomic resonance Faraday filter include its wide field of view, high background rejection and high peak transmission (Dick and Shay 1991). A Faraday filter is constructed by placing a circularly birefringent, dichroic medium between crossed polarizers, as shown in Figure 4.1. The medium can be an atomic vapor in an axial magnetic field, causing a rotation of the polarization of light near an atomic resonance while the polarization of off-resonance light is unaffected. This is a result of the interaction of the electromagnetic light wave with the atomic electronic states.

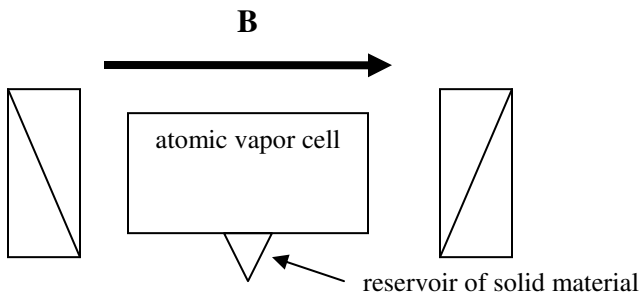

Figure 4.1: Schematic of an atomic vapor Faraday filter, consisting of a vapor cell in an axial magnetic field between crossed polarizers.

The Faraday effect was first observed near atomic resonances by Macaluso and Corbino (1898), and a Faraday filter of this type was first introduced by Ohman (1956). Studies include using the Rb 5s-5p (Dick and Shay 1991), the Cs 6s-6p (Menders, et al. 1991), and Ca 4s-4p transitions (Chan and Gelbwachs 1993). In 1975, Agnelli et al. (1975) constructed a Na Faraday filter with an axial magnetic field of 1500 G for observation of solar Na D lines. This filter had a 0.016 nm full-width and a peak transmission of 25%. Studies of the relationship between vapor temperature and cell transmission were performed by Hu et al. (1998) and Zhang et al. (2001) for Na and K filters respectively. In 1993, Chen et al. (1993) developed a Na filter with a 0.002 nm full-width and a peak transmission of 85% (excluding passive losses from optical components), and a background fractional transmission of 2×10^{-5}. This was used for daytime Na fluorescence lidar temperature measurements (Chen, et al. 1996). Further developments have led to a Faraday filter deployed in the Colorado State University Na lidar system, allowing the measurement of mesopause region temperature and horizontal wind under sunlit conditions, thus permitting studies of the solar atmospheric tides and their variability (She, et al. 2004).

Designing the FFBS for this study requires theoretical calculations of filter transmissions in order to optimize instrument parameters. This chapter details the complete theoretical calculations of Faraday filter transmission beginning with first principles. A summary of this work was also presented in Harrell et al. (2009).

In Section 4.1, I will detail the classical calculation of the filter transmission and then relate this to the quantum-mechanical derivation in Section 4.2. The connection between the classical electric dipole and quantum mechanical matrix elements will also

be shown. Section 4.3 explains the theoretical results, where I have also tabulated current values for all relevant constants and coefficients.

4.1 Classical theory of filter transmission

The transmission of the filter as a function of frequency can be derived using the classical theory of electricity and magnetism along with a simple model for the response of sodium atoms. First, the atom is assumed to consist of an electron with charge e and reduced mass m_e bound to the nucleus as a classical harmonic oscillator with resonant angular frequency ω_0 acted on by external electric and magnetic fields **E** and **B**. The equation of motion can be written:

$$m_e\ddot{\mathbf{x}} + m_e\Gamma\dot{\mathbf{x}} + m_e\omega_0^2\mathbf{x} = -e\mathbf{E} - e\dot{\mathbf{x}}\times\mathbf{B}, \tag{4.1}$$

where Γ is a phenomenological damping constant to be defined later (Jackson 1999). The oscillating electric field is due to the incident light, which we will take as traveling along the z-axis. In the electric dipole approximation, the force from the oscillating magnetic field is negligible, and so the only the constant applied axial field along the z-axis, B_0 is included. Solving for x- and y- components yields:

$$\begin{aligned} m_e x(\omega^2 + i\Gamma\omega - \omega_0^2) - ei\omega y B_0 &= eE_x, \\ m_e y(\omega^2 + i\Gamma\omega - \omega_0^2) + ei\omega x B_0 &= eE_y. \end{aligned} \tag{4.2}$$

Later it will be convenient to consider circular polarizations as they are eigenstates for the vapor medium under a constant axial magnetic field. Circular polarization is just a superposition of x- and y- components of equal magnitude, oscillating 90 degrees out of phase, so for the left and right circular polarizations,

$$\mp\frac{(x\pm iy)}{\sqrt{2}}\left[m_e(\omega^2+i\Gamma\omega-\omega_0^2)\mp e\omega B_0\right]=\mp\frac{e\left(E_x\pm iE_y\right)}{\sqrt{2}}, \tag{4.3}$$

with $\hat{x}$ and $\hat{y}$ coordinates written in terms of circular $\hat{+}$ and $\hat{-}$ coordinates:

$$\hat{x}=-\frac{\hat{+}-\hat{-}}{\sqrt{2}}\text{, and }\hat{y}=i\frac{\hat{+}+\hat{-}}{\sqrt{2}}. \tag{4.4}$$

To calculate filter transmission, we express the optical response of the medium in terms of a complex susceptibility $\chi_\pm(\omega)=\chi'_\pm+i\chi''_\pm$, where the ± represents left- and right-handed circular polarizations of the incident light. The complex susceptibility relates the wavenumber k and angular frequency ω in the dispersion relation as

$$k=\frac{\omega}{c}\sqrt{(1+\chi'_\pm+i\chi''_\pm)}\approx\frac{\omega}{c}\left(1+\frac{\chi'_\pm}{2}+i\frac{\chi''}{2}\right), \tag{4.5}$$

where c is the speed of light. This is valid for relatively low vapor density, a situation suitable for most applications. The quantum mechanical derivation of $\chi_\pm(\omega)$ is given in Section 4.2.

The electric field of electromagnetic radiation may be written as

$$\mathbf{E}=E_0\exp\left(i(kz-\omega t)\right)\hat{\mathbf{x}}=E_0\exp\left(i\frac{\omega}{c}\left(1+\frac{\chi'_\pm}{2}+\frac{i\chi''_\pm}{2}\right)z-\omega t\right)\hat{\mathbf{x}}. \tag{4.6}$$

After going through the first polarizer of the Faraday filter, the randomly polarized input electric field is linearly polarized, which can be expressed as a sum of counter-rotating circular polarizations.

The input electric field, $\mathbf{E}(z,t)=\frac{1}{2}\left[\mathcal{E}(z,\omega)e^{-i\omega t}+\mathcal{E}^*(z,\omega)e^{i\omega t}\right]$, enters the vapor cell at z = 0 and t=0. After passing through the vapor cell with length z=L and time t, the light the electric field will be:

$$\mathcal{E}(L,\omega) = -\frac{E_0}{\sqrt{2}}\left[\exp\left(i\left\{\frac{\omega}{c}(1+0.5\chi'_+ + 0.5i\chi''_+)L\right\}\right)\hat{+}\right.$$
$$\left.-\exp\left(i\left\{\frac{\omega}{c}(1+0.5\chi'_- + 0.5i\chi''_-)L\right\}\right)\hat{-}\right]. \tag{4.7}$$

which contains (for each circular polarization) an exponential decay or absorption term depending on χ'', and an oscillatory term depending on $1+0.5\chi'$—this is the index of refraction.

A second, crossed polarizer will select the light polarized in the $\hat{y}$-direction, so the transmission coefficient will be:

$$\mathcal{T}(\nu) = \frac{\left|\mathcal{E}\bullet\hat{y}\right|^2}{E_0^2} = \frac{1}{4}\left[\exp\left(-\frac{\omega}{c}\chi''_+ L\right) + \exp\left(-\frac{\omega}{c}\chi''_- L\right)\right.$$
$$\left.-2\exp\left(-\frac{\omega}{c}\frac{\chi''_+ + \chi''_-}{2}L\right)\cos\left(\frac{\omega}{c}\frac{\chi'_+ - \chi'_-}{2}L\right)\right]. \tag{4.8}$$

We can define the Faraday rotation θ_F as the angle of polarization rotation of the output light relative to the initial linear polarization:

$$\theta_F = \frac{\omega}{c}\frac{\chi'_+ - \chi'_-}{4}L = \frac{\pi}{2\lambda}(\chi'_+ - \chi'_-)L = \frac{\pi}{\lambda}(\Delta n)L, \tag{4.9}$$

where Δn is the difference in the index of refraction of the two circular polarizations (i.e., circular birefringence).

While this classical calculation provides an excellent understanding of the transmission of light through the cell in terms of susceptibilities, it is unfortunately an incomplete description if we wish to produce valid numerical values for design or application purposes. To obtain numerical results requires a quantum mechanical calculation of the susceptibilities.

4.2 Quantum-mechanical theory

4.2.1 Derivation of susceptibility

Since an atomic vapor is an ensemble of many atoms, its state may be represented by a density matrix: $\rho = \sum_n |\psi_n\rangle p_n \langle\psi_n|$ (Sakurai 1994). To model the interaction of the atom in an external magnetic field, perturbation theory is used. The Schrödinger equation for the evolution of the density matrix of a system with Hamiltonian $H_0+\gamma H^I$, including a damping term to model interactions such as collisions, is:

$$\begin{aligned}\frac{d\rho_{\alpha\beta}(t)}{dt} &= \frac{1}{i\hbar}\left\{[H_0,\ \rho(t)]_{\alpha\beta} + [\gamma H^I(t),\ \rho(t)]_{\alpha\beta}\right\} + \left.\frac{\partial\rho_{\alpha\beta}}{\partial t}\right|_{random} \\ &= \frac{1}{i\hbar}\left\{\hbar\omega_{\alpha\beta}\rho_{\alpha\beta} + [\gamma H^I(t),\ \rho(t)]_{\alpha\beta}\right\} - \Gamma_{\alpha\beta}\rho_{\alpha\beta}(t),\end{aligned} \tag{4.10}$$

where $\rho_{\alpha\beta}(t)$ represents an element of the density matrix between energy eigenstates of H_0 α and β, where $\omega_{\alpha\beta}$ is the associated transition frequency. The derivative $\left.\frac{\partial\rho_{\alpha\beta}}{\partial t}\right|_{random}$ represents the effect of random collisions between the atoms. The damping constant $\Gamma_{\alpha\beta}$ is related to the natural linewidth of the transition. This means that $\Gamma_{\alpha\beta}/\pi = A_{\alpha\beta}/(2\pi)$, where $A_{\alpha\beta}$ is the Einstein coefficient for the transition rate between the two states. H_0 is the base, or unperturbed, Hamiltonian, and $H^I(t)$ is the interaction Hamiltonian, with γ the strength of the interaction.

In perturbation theory $\rho_{\alpha\beta}$ can be expressed as a power series: $\rho_{\alpha\beta}(t) = \sum_{i=0}^{\infty}\gamma^i\rho_{\alpha\beta}^{(i)}$ with each term of decreasing importance obtained from a hierarchy of equations. Since we are interested in a linear susceptibility, only the first-order equation for determining $\rho_{\alpha\beta}^{(1)}$ is needed:

$$\frac{d\rho_{\alpha\beta}^{(1)}(t)}{dt} = \frac{1}{i\hbar}\left\{\hbar\omega_{\alpha\beta}\rho_{\alpha\beta}^{(1)}(t) + \left[H^{I}(t), \rho^{(0)}\right]_{\alpha\beta}\right\} - \Gamma_{\alpha\beta}\rho_{\alpha\beta}^{(1)}(t). \tag{4.11}$$

For the electric dipole approximation,

$H^{I}(t) = -e\mathbf{r}\bullet\mathbf{E}(t) = -\frac{e}{2}\sum_{k} r^{k}\left[\mathcal{E}^{k}(\omega)e^{-i\omega t} + \mathcal{E}^{*k}(\omega)e^{i\omega t}\right]$, with $-er_{\alpha\beta}^{k} = -e\langle\alpha|r^{k}|\beta\rangle$ and

k=x,y as the x- and y-components of the electric dipole moment of an atomic electron with charge e connecting the |α> and |β> eigenstates, and $\mathbf{E}(t)$ is the electric field of light propagating along the axis of the Faraday filter. Defining

$$\rho_{\alpha\beta}^{(1)}(t) = \frac{1}{2}\left[\rho_{\alpha\beta}^{(1)}(\omega)e^{-i\omega t} + \rho_{\alpha\beta}^{(1)^{*}}(\omega)e^{i\omega t}\right], \tag{4.12}$$

the steady-state solution to Eq. (4.11) is:

$$\rho_{\alpha\beta}^{(1)}(\omega) = \frac{1}{\hbar}\sum_{k=x,y}\frac{e\left(r_{\alpha\beta}^{k}\rho_{\beta\beta}^{(0)} - \rho_{\alpha\alpha}^{(0)}r_{\alpha\beta}^{k}\right)\mathcal{E}^{k}}{\omega - \omega_{\alpha\beta} + i\Gamma_{\alpha\beta}}. \tag{4.13}$$

The expectation value of the polarization is

$$\mathbf{P}(t) = \langle -Ne\mathbf{r}\rangle = -Ne\,\mathrm{Tr}\left[\rho^{(1)}(t)\mathbf{r}\right] = \frac{1}{2}\left[\mathcal{P}(\omega)e^{-i\omega t} + \mathcal{P}^{*}(\omega)e^{+i\omega t}\right], \tag{4.14}$$

where N is the vapor density.

To avoid confusion in notation, we decompose the vectors, $\mathcal{P}(\omega)$, $\mathbf{E}(\omega)$ and $\mathbf{r}$ into Cartesian coordinates. For an isotropic medium, both the polarization and electric field are transverse to the propagation direction $\hat{z}$, so combining (4.13) and (4.14) gives:

$$\mathcal{P}_{j}(\omega) = -\frac{Ne^{2}}{\hbar}\sum_{\alpha,\beta,k}\frac{\left(\rho_{\beta\beta}^{(0)} - \rho_{\alpha\alpha}^{(0)}\right)r_{\beta\alpha}^{j}r_{\alpha\beta}^{k}E^{k}(\omega)}{\omega - \omega_{\alpha\beta} + i\Gamma_{\alpha\beta}}. \tag{4.15}$$

where j and k represent the components $(\hat{x}, \hat{y})$.

Due to the axial symmetry of the magnetic field, the circular polarizations are eigenmodes of the system. We can transform the Cartesian coordinates into circular polarizations using (4.4), and rewrite (4.15) in terms of the relevant dipole moment, $-er^{\pm}_{\alpha\beta} = -e\langle\alpha|r^{\pm}|\beta\rangle$, transition frequency, $\omega^{\pm}_{\alpha\beta}$, and damping constant, $\Gamma^{\pm}_{\alpha\beta}$. We then have:

$$\mathcal{P}_{\pm}(\omega) \equiv \varepsilon_0 \chi^{(1)}_{\pm}(\omega)\mathcal{E}_{\pm}(\omega) \rightarrow \chi^{(1)}_{\pm}(\omega) = -\frac{Ne^2}{\varepsilon_0 \hbar}\sum_{\alpha\beta}\frac{\left(\rho^{(0)}_{\beta\beta}\left|\left(r^{\pm}\right)_{\beta\alpha}\right|^2 - \rho^{(0)}_{\alpha\alpha}\left|\left(r^{\pm}\right)_{\alpha\beta}\right|^2\right)}{\omega - \omega^{\pm}_{\alpha\beta} + i\Gamma^{\pm}_{\alpha\beta}}. \tag{4.16}$$

By assuming that in the zero-order, only the ground state is populated, we can rewrite Eq. (4.16) (replacing β with *g* for "ground state"):

$$\chi^{(1)}_{\pm}(\omega) = -\frac{Ne^2}{\varepsilon_0 \hbar}\sum_{\alpha g}\rho^{(0)}_{g}\left|r^{\pm}_{\alpha g}\right|^2\left(\frac{1}{\omega - \omega^{\pm}_{\alpha g} + i\Gamma^{\pm}_{\alpha g}} - \frac{1}{\omega + \omega^{\pm}_{\alpha g} + i\Gamma^{\pm}_{\alpha g}}\right), \tag{4.17}$$

where $\omega^{\pm}_{\alpha g} = -\omega^{\pm}_{g\alpha}$ and $\Gamma^{\pm}_{\alpha g} = \Gamma^{\pm}_{g\alpha}$ have been assumed.

Since the resonance line is narrow $\left(\Gamma^{\pm}_{\alpha g} << \omega\right)$, and for the range of frequencies of interest, $\omega + \omega^{\pm}_{\alpha g} \approx 2\omega^{\pm}_{\alpha g} >> \Gamma^{\pm}_{\alpha g}$, Eq. (4.17) reduces to

$$\chi^{(1)}_{\pm}(\omega) = \frac{Ne^2}{\varepsilon_0 \hbar}\sum_{\alpha g}\rho^{(0)}_{g}\left|\left(r^{\pm}\right)_{\alpha g}\right|^2\left(\frac{2\omega^{\pm}_{\alpha g}}{2\omega^{\pm}_{\alpha g}\left(\omega^{\pm}_{\alpha g} - \omega - i\Gamma^{\pm}_{\alpha g}\right)}\right). \tag{4.18}$$

Including a Doppler broadening due to an atom's random motion with Gaussian distributed line of sight speed v, equation (4.18) becomes:

$$\chi^{(1)}_{\pm}(\nu) = \frac{N}{2\pi\varepsilon_0 \hbar}\frac{1}{\sqrt{\pi}u}\sum_{\alpha g}\rho^{(0)}_{\alpha g}\int_{-\infty}^{\infty}\frac{\left|\left(p_{\pm}\right)_{\alpha g}\right|^2 \exp\left(-\mathrm{v}^2/u^2\right)d\mathrm{v}}{\left[\nu^{\pm}_{\alpha g} - (\mathrm{v}/\lambda) - \nu - i\left(\Gamma^{\pm}_{\alpha g}/2\pi\right)\right]}, \tag{4.19}$$

where $u = \sqrt{2k_B T / m}$, and k_B, T, and m are respectively, the Boltzmann constant, temperature and atomic mass. We have also replaced $e^2 \left| r_{\alpha g}^{\pm} \right|^2$ with $\left| \left(p_{\pm} \right)_{\alpha g} \right|^2$ for electric dipole moment. The integral in (4.19) is the complex Faddeeva function (Schreier 1992), and can be rewritten in real and imaginary components as:

$$\chi_{\pm}^{'}(\nu) = \frac{N}{2\pi\varepsilon_0 \hbar} \frac{1}{\sqrt{\pi} u} \sum_{\alpha g} \rho_{\alpha g}^{(0)} \int_{-\infty}^{\infty} \frac{\left| \left(p_{\pm} \right)_{\alpha g} \right|^2 \left(\nu_{\alpha g}^{\pm} - (\nu/\lambda) - \nu \right) \exp\left(- \nu^2 / u^2 \right) d\nu}{\left[\left(\nu_{\alpha g}^{\pm} - (\nu/\lambda) - \nu \right)^2 + \left(\Gamma_{\alpha g}^{\pm} / 2\pi \right)^2 \right]}, \quad (4.20)$$

$$\chi_{\pm}^{''}(\nu) = \frac{N}{2\pi\varepsilon_0 \hbar} \frac{1}{\sqrt{\pi} u} \sum_{\alpha g} \rho_{\alpha g}^{(0)} \int_{-\infty}^{\infty} \frac{\left| \left(p_{\pm \alpha g} \right) \right|^2 \left(\Gamma_{\alpha g}^{\pm} / 2\pi \right) \exp\left(- \nu^2 / u^2 \right) d\nu}{\left[\left(\nu_{\alpha g}^{\pm} - (\nu/\lambda) - \nu \right)^2 + \left(\Gamma_{\alpha g}^{\pm} / 2\pi \right)^2 \right]}. \quad (4.21)$$

In practice, the ground state is a multiplet, and the factor $\rho_g^{(0)}$ takes into account any differences in fractional population in the various ground states in equilibrium with each other at temperature T. From Maxwell-Boltzmann statistics $\rho_g^{(0)}$ is:

$$Z = \sum_{i=1}^{N} \exp\left(- \frac{\mathsf{E}_i}{k_B T} \right) \rightarrow \rho_g{}^{(0)} = \frac{N_i}{N_g} = \frac{\exp\left(- \frac{\mathsf{E}_i}{k_B T} \right)}{Z}, \quad (4.22)$$

where E_i, not to be confused with the electric field, is the ground-state energy eigenvalue for an individual state from the Appendix.

4.2.2 Derivation of Transmission matrix elements

The transition matrix elements depend on the atomic energy structure, which is shown in Fig. 4.2 for the 3s and 3p states in a Na atom, in the absence of an external magnetic field. The Hamiltonian governing the splitting of individual energy levels due to the hyperfine structure and Zeeman splitting is (Corney 1977):

$$
\begin{aligned}
H &= H_0 + H^I \\
H^I &= H_{HFS} + H_{Zeeman} \\
&= A_J(\mathbf{I}\bullet\mathbf{J}) + \frac{B_J}{2I(2I-1)J(2J+1)}\left[3(\mathbf{I}\bullet\mathbf{J})^2 + \tfrac{3}{2}(\mathbf{I}\bullet\mathbf{J}) - I(I+1)J(J+1)\right] \\
&\quad + g_J\mu_B B_0\mathbf{J} - g_I\mu_N B_0\mathbf{I},
\end{aligned}
\tag{4.23}
$$

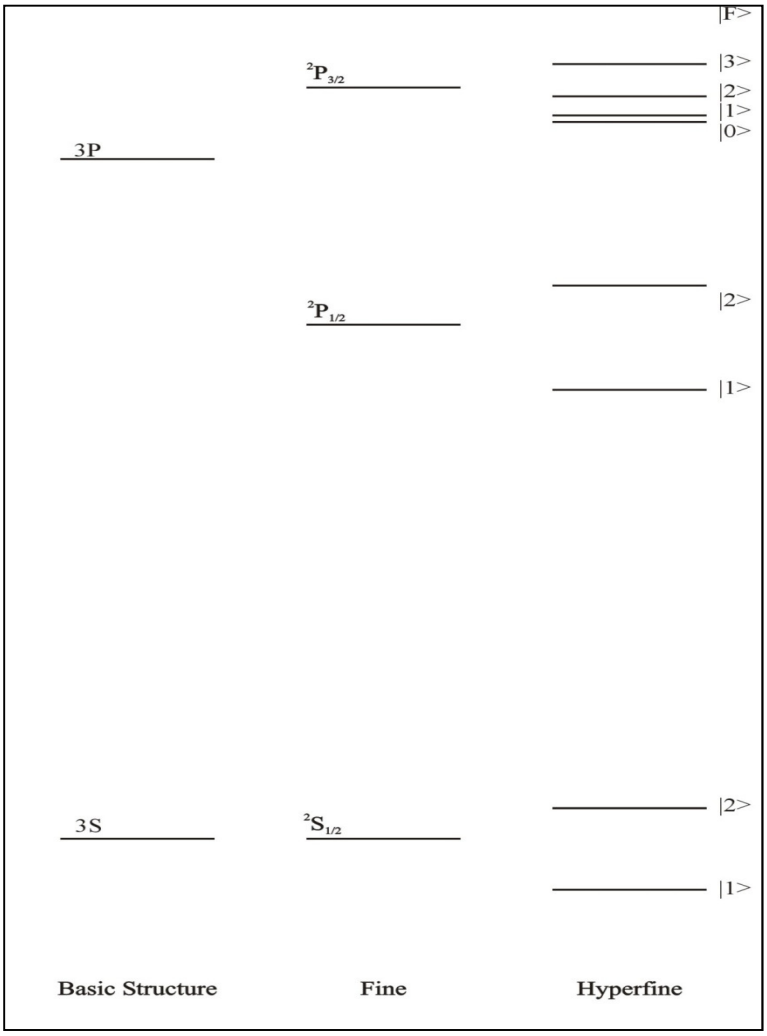

Figure 4.2: Basic, Fine and Hyperfine structure for 3S and 3P states of the Na atom (not to scale).

where H_0 and H^I are the base and interaction Hamiltonians, respectively. The parameters A_J and B_J represent the strength of the hyperfine magnetic dipole and electric quadrupole interactions; vectors **I** and **J** are the nucleus and electron total angular momentum operators with magnitudes I and J; μ_B and μ_N are the Bohr and nuclear magnetons (1.399×10^{-3} GHz G^{-1} and 7.623×10^{-7} GHz G^{-1}, respectively); B_0 is the external magnetic field strength; and g_J and g_I are the Lande-g factors of the atom and of the nucleus; g_J may be related to J, orbital angular momentum L and electron spin S as (Corney 1977):

$$g_J = \frac{3J(J+1) + S(S+1) - L(L+1)}{2J(J+1)} . \qquad (4.24)$$

In order to evaluate the transition matrix element of the dipole moment in (4.19) and the energies of the eigenstates, we must first solve the eigenvalue problem of the individual atomic states. The first step is to choose a set of basis states; here the base Hamiltonian includes the Coulomb attraction of the nucleus, the interactions between atomic electrons, and the spin-orbit interaction. Thus the eigenstates of H_0 are $|\gamma\ L\ S\ J\rangle$ with total electronic angular momentum J and nuclear spin I as good quantum numbers for H^I, and they have a degeneracy of $(2J+1)(2I+1)$ (Corney 1977). Under the influence of H^I, some of the degeneracy will be lifted. In the limit of zero (or low) external magnetic field, the coupling between the atomic electrons and the nucleus will be dominant, and the eigenstates of $H_0 + H_{HFS}$ are $|I\ J\ F\ m_F\rangle$, with the total angular momentum of the atom (including the nucleus) **F=I+J**; the energy eigenvalues will then depend of the strengths of hyperfine interactions, A_J and B_J. In the limit of high external magnetic field, its interaction with the atom will cause **I** and **J** to align to the external field separately. This would allow the use of $|I\ J\ m_I\ m_J\rangle$ as the eigenstates of $H_0 + H_{Zeeman}$; however to account for hyperfine interactions, the eigenvalues include the

diagonal contribution from H_{HFS} as an approximation. For a solution valid at all values of magnetic field strength, an appropriate choice would be $|I\ J\ Q\ m_Q>$, which reduces to $|I\ J\ F\ m_F>$ in the low or zero field limit, and to $|I\ J\ m_I\ m_J>$ in the high field limit. I choose to write the $|J\ I\ Q\ m_Q>$ states as a linear superposition of $|I\ J\ m_I\ m_J>$ states (Corney 1977). These eigenstates of H_0+H_{Zeeman} can be abbreviated $|m_I\ m_J>$. The total Hamiltonian in this representation is not diagonal; it must then be diagonalized to obtain the eigenstates and eigenvalues of the system. Figure 4.3 shows an example of splitting in the zero field limit (middle column) and the high field limit (right column) for the sodium $^2P_{3/2}$ excited state.

The solutions for energy eigenvalues, eigenvectors, and transition matrix elements are in the Appendix. With the electric dipole matrix transition elements $\left|(p_{\pm})_{\alpha g}\right|$ evaluated, Eqs. (4.20) and (4.21) may be used to calculate χ' and χ'', from which $\mathfrak{I}(\nu)$ and θ_F may be calculated via (4.8) and (4.9).

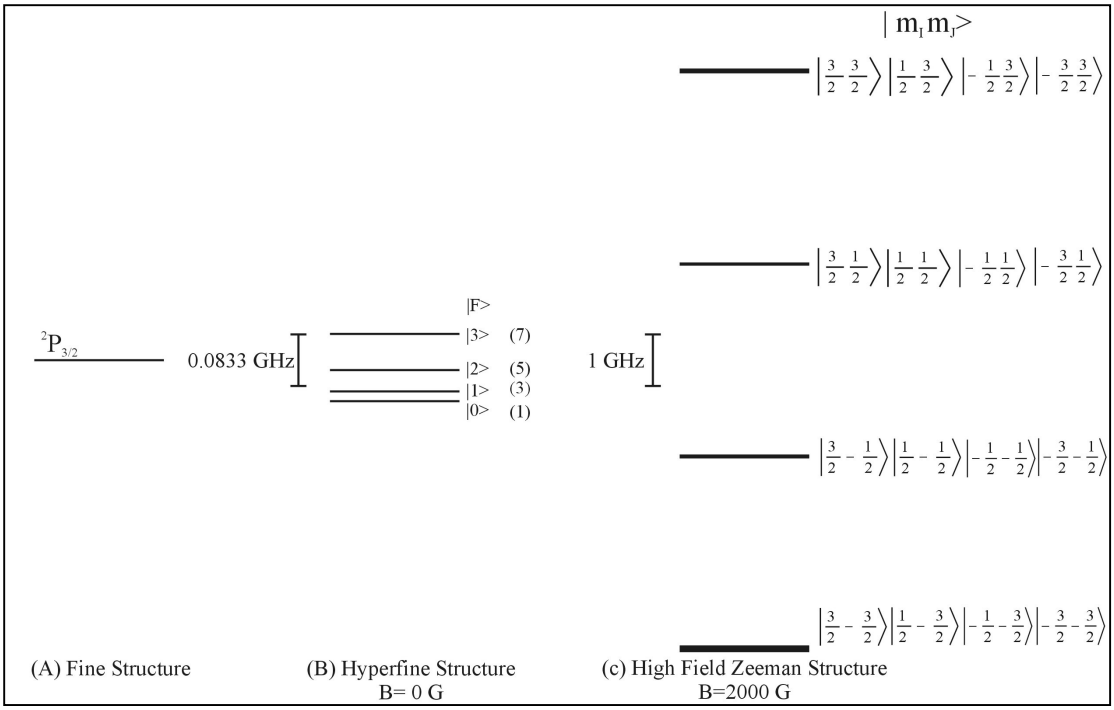

Figure 4.3: Energy level diagram for Na showing the $3^2P_{3/2}$ excited states. (a) shows the single fine structure state, with a degeneracy of 16. (b) is the hyperfine splitting case with no external magnetic field. The $|F\rangle$ eigenstate notation and degeneracies (in parentheses) are indicated to the right and the scale in GHz is to the left. (c) shows the exact solution for the Zeeman structure for an external magnetic field of 2000G. The states to be broken up into 4 closely spaced groups due to the different values for μ_I and μ_B. Each Zeeman state has a degeneracy of 1. The $|m_I\, m_J\rangle$ eigenstate notations are listed to the right and the scale in GHz is to the left. The numerical values of the splitting are listed in Table 4.1 and Table 4.2.

Table 4.1: Hyperfine structure state offset from fine structure (GHz).

$\|F\rangle$	$^2P_{3/2}$
|3>	0.0424678
|2>	-0.0159712
|1>	-0.0503922
|0>	-0.0662413

Table 4.2: Zeeman effect state offset (GHz) from fine structure for B_0=2000 G.

$\|m_I\ m_J\rangle$	$^2P_{3/2}$
$\left\|\frac{3}{2}\frac{3}{2}\right\rangle$	5.63525
$\left\|\frac{3}{2}\frac{1}{2}\right\rangle$	1.87503
$\left\|\frac{1}{2}\frac{3}{2}\right\rangle$	5.60852
$\left\|\frac{3}{2}\frac{-1}{2}\right\rangle$	-1.88366
$\left\|\frac{1}{2}\frac{1}{2}\right\rangle$	1.86958
$\left\|\frac{-1}{2}\frac{3}{2}\right\rangle$	5.58296
$\left\|\frac{3}{2}-\frac{3}{2}\right\rangle$	-5.64084
$\left\|\frac{1}{2}-\frac{1}{2}\right\rangle$	-1.87066
$\left\|-\frac{1}{2}-\frac{1}{2}\right\rangle$	1.86273
$\left\|-\frac{3}{2}\frac{3}{2}\right\rangle$	5.55863
$\left\|\frac{1}{2}-\frac{3}{2}\right\rangle$	-5.61218
$\left\|-\frac{1}{2}-\frac{1}{2}\right\rangle$	-1.85894
$\left\|-\frac{3}{2}\frac{1}{2}\right\rangle$	1.85444
$\left\|-\frac{1}{2}-\frac{3}{2}\right\rangle$	-5.58202
$\left\|-\frac{3}{2}-\frac{1}{2}\right\rangle$	-1.84853
$\left\|-\frac{3}{2}-\frac{3}{2}\right\rangle$	-5.55032

4.3 Computer program results and discussion

A computer program was written in the IDL language (ITT Corporation), using built-in functions to calculate eigenvalues and eigenstates, which agreed exactly with longhand calculations. The complex Faddeeva function was calculated using an algorithm reported by Schreier (1992).

Filter transmissions and associated θ_F were calculated for both D_2 and D_1 transitions. Tables 4.3 and 4.4 give the necessary constants and coefficients. Values of g_I were calculated using the method detailed in (Krane 1988).

One parameter is yet to be defined: vapor density N. The density of the atomic vapor is a function of its temperature and pressure. In our particular vapor cell, we control the temperature of the reservoir (see Fig. 4.1) containing solid or liquid Na to set the vapor pressure on the coexistence curve. Another controller fixes the temperature of the main body of the cell which is the vapor temperature. From the ideal gas law the vapor density, which is proportional to this vapor pressure and inversely proportional to the temperature of the vapor, can be determined. Data for vapor pressure as a function of saturated vapor temperature is given for both solid or liquid Na by Honig and Kramer (1969), and an equation was derived using a curve-fit method similar to that described in that paper. For Na, the vapor pressure equation is:

$$\begin{aligned}\log_{10}(P_{Na}) &= 71.899 - 9217.2(T_{res})^{-1} + 40693000(T_{res})^{-3} \\ &+ 0.0061264(T_{res}) - 9.6625\ln(T_{res})\end{aligned}, \qquad (4.25)$$

where T_{res} is the temperature at the solid or liquid reservoir, and P is the pressure in Torr.

Figures 4.4 and 4.5 show transmission functions and susceptibilities calculated from the computer program, with the necessary values and filter parameters listed in Tables 4.3 through 4.5. In the curves for χ' and χ'' shown in Fig. 4.5, there are two groups of curves. Each set of curves is a summation of the curves for the various Zeeman split transitions with $\Delta m_I = \pm 1$, which is the selection rule for the circular $\hat{+}$ and $\hat{-}$ polarizations defined in Section 4.1. The central value of each group, indicated by the peak in the curve for χ'' or the zero crossing point of χ' is labeled as ν_0^- for the curves on the left and ν_0^+ for the curves on the right, and can be thought of as a resonance frequency for the circular polarization. This allows χ to be written as $\chi_\pm$ as was done in Eqs. (4.20) and (4.21).

By comparing the figures with the relevant equations, we can gain physical insight to what occurs in a Faraday filter. To facilitate this appreciation, we further consider the transmission function near three specific frequencies: A. at line center, B. off resonance, and C. at the resonance at either ν_0^+ or ν_0^- (see figure 4.4a). For cases A and B, there is no absorption and the exponential factors in Eq. (4.8) are all unity, and the transmission reduces to $\mathcal{T}(\nu) = \frac{1}{2}[1 - \cos(2\theta_F)] = \sin^2\theta_F$. For case A, the θ_F should be nonzero, since χ_+' and χ_-' are of opposite sign and $\mathcal{T}(\nu)$ varies periodically between 0 and near 1 as vapor density varies, so for fixed B_0 and cell length, we can easily adjust the vapor density to vary transmission at the line center. For case B, $\theta_F = 0$ since far from resonance both χ_+' and χ_-' have the same value and $\mathcal{T}(\nu) = 0$. For case C, at ν_0^+ for example, χ_+' is zero; while χ_-'' is non-zero, χ_+'' is large enough to render the associated exponential factors zero. In this case, one circular polarization is totally absorbed, and the electric field of the other polarization is reduced by the second polarizer by a factor of $\frac{1}{\sqrt{2}}$. Therefore, $\mathcal{T}(\nu) \cong 0.25$ at either ν_0^+ or ν_0^-. In practice, the measured transmission can be scaled to ¼ at those frequencies to avoid the more involved normalization measurement described by Chen, et al. (1993).

Table 4.3: Necessary values for calculation.

	Na		
m (u)	22.989768		
g_I	1.478392		
%	100%		
	A_J (GHz)	B_J (GHz)	g_J
$^2S_{1/2}$	0.8858130644	0	2
$^2P_{1/2}$	0.0944	0	2/3
$^2P_{3/2}$	0.018572	0.002723	4/3

Table 4.4: Linestrength*, transition vacuum wavelength and linewidth.

	Na	
	D_1	D_2
λ (nm)	589.7558	589.1582
$S=2S_0$ ($e^2a_0^2$)	37.3	37.3
Δν (GHz)	0.00977	0.00980

*Linestrength S_0 is defined in the Appendix.

Table 4.5: Filter Parameters used to generate Figs. 4.4-4.5, and transition frequencies

	Na	
ℓ (cm)	4	
B_0 (G)	1850	
	D_1	D_2
T_{res} (°C)	180	164
T_{cell} (°C)	186	169
$D_1\nu_0^+$ (GHz)	3.6	
$D_1\nu_0^-$ (GHz)	-3.6	
$D_2\nu_0^+$ (GHz)	2.5	
$D_2\nu_0^-$ (GHz)	-2.5	

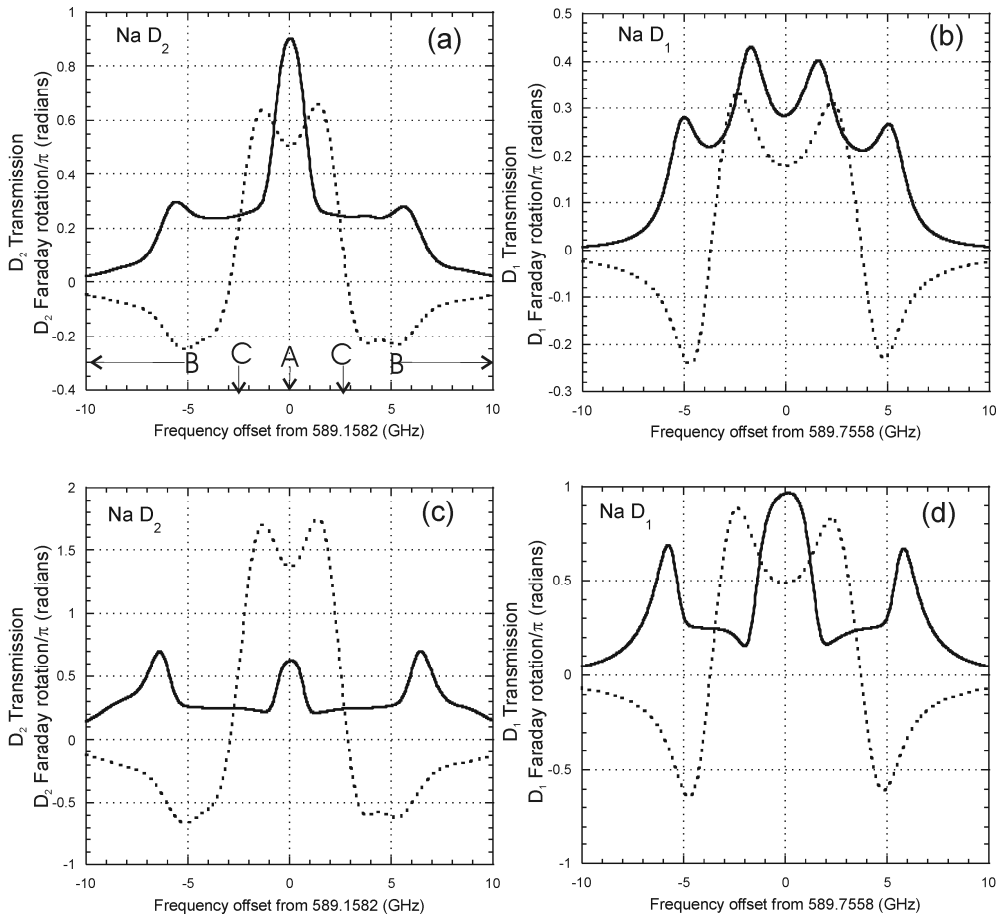

Figure 4.4: D_2 (a) and D_1 (b) transmission as a fraction of input linear polarization (solid) and Faraday rotation in units of π (dashed) vs. frequency (GHz) for a Na vapor Faraday filter optimized for maximium transmission at the D_2 line center. (c) and (d) are the same for a D_1 optimized filter. Filter parameters are listed in Table 4.5.

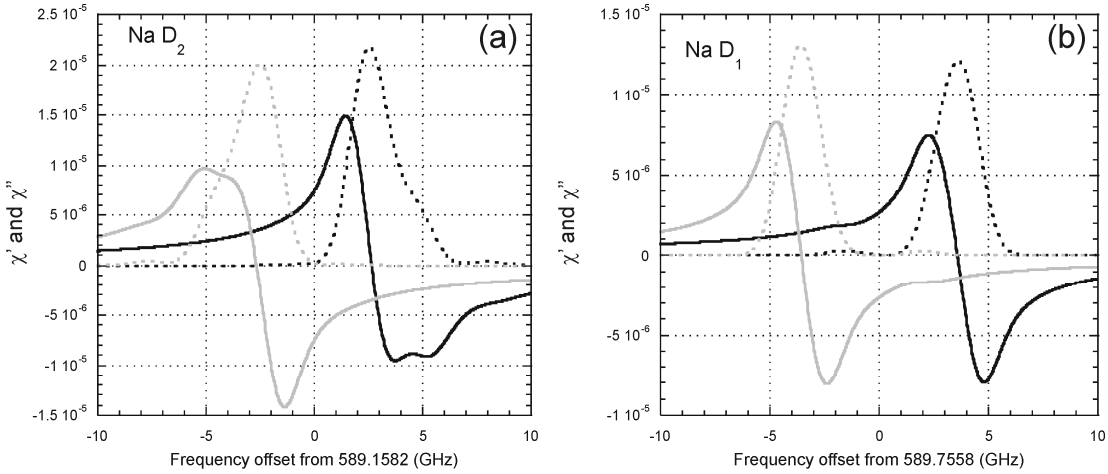

Figure 4.5: χ' and χ'' curves for Na D_2 (a) and D_1 (b) lines split by the Zeeman effect due to the 1850 Gauss external magnetic field for the D_2 optimized filter. Solid gray curve is χ' for σ_-, gray dashed is χ'' for σ_-, black solid is χ' for σ_+, and black dashed is χ'' for σ_+.

4.4. Conclusions: Other atmospheric studies using Faraday filters

Several other instruments of note have been designed to study the atmosphere utilizing Faraday filters. The aforementioned sodium lidar receiver is perhaps the most significant, as it allows 24-hour continuous measurements of temperature and zonal and meridional wind, culminating in numerous studies of mesopause region tidal waves (Yuan, et al. 2008; Yuan, et al. 2006). Work on potassium Faraday filters for use in the receiver of a potassium lidar has also seen success (Fricke-Begemann 2004). Also significant is the recent work by Huang et al. to use Faraday filters as a double-edge filter in the receiver of a 3-frequency sodium lidar to measure lower atmosphere temperature and winds (Huang, et al. 2009). Yet a third use is the Magnetooptic Doppler Analyzer developed by Williams for determining mesopause region winds from the sodium nightglow (Williams and Tomczyk 1996).

It is worth noting that while some of these same results could have been achieved using other filters (such as Fabry-Perot interferometers in lidar receivers, see Hoffner and

Lautenbach 2009), the Faraday filters described here have many advantages, such as lower cost, ease of operation and extremely narrow bandwidth unachievable by other methods.

CHAPTER 5

THE FARADAY FILTER-BASED SPECTROMETER

The Faraday Filter-Based spectrometer (FFBS) is a device designed to measure R_{D1}, R_{D2}, R_{DX}, R_{DA}, and R_D of the Na nightglow. Chapters 2 and 3 detailed why this measurement is important and how the FFBS can achieve it; the Faraday filter description in Chapter 4 describes how the narrow filter bandwidths needed to measure R_{D1}, R_{D2}, R_{DX}, and R_{DA} are achieved. In this chapter, I will describe the various other components of the optical layout, the electronic control and data collection system, the method for choosing Faraday filter temperature settings, and the data analysis procedure.

5.1 FFBS optical layout

The optical layout of the FFBS is shown in figure 5.1, and photographically in figure 5.2. The nightglow enters the vertically-pointing Celestron 14-inch Schmidt-Cassegrain telescope or Starsplitter 30-inch Newtonian telescope, and then is fiber-coupled (much like the CSU Na lidar to be described in Chapter 6) into the spectrometer. At the input to the spectrometer, the light is then roughly collimated with a 25 mm focal length achromatic doublet lens. It then passes through a filter wheel, where the light passes through one of two ~0.3 nm FWHM interference filters: one is centered at the D_2 wavelength (589.158 nm), while the other is centered at the D_1 wavelength (589.756 nm).

Then, the unpolarized light is split into its two orthogonal polarizations by the Glan-Taylor polarizing beamsplitter, which has

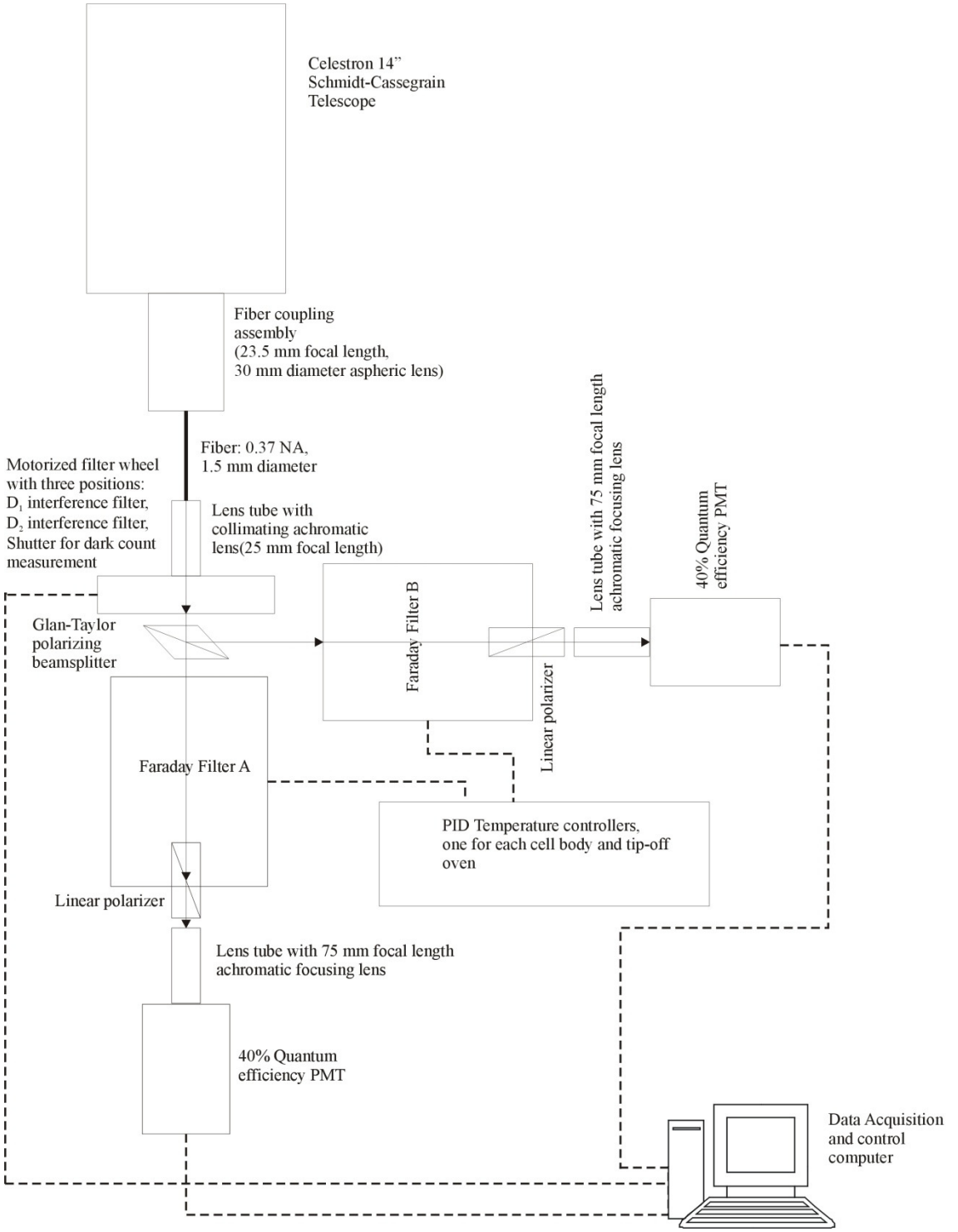

Figure 5.1: Layout of the FFBS. Solid lines with arrowheads indicate the light path, while dotted lines indicate electronic connections.

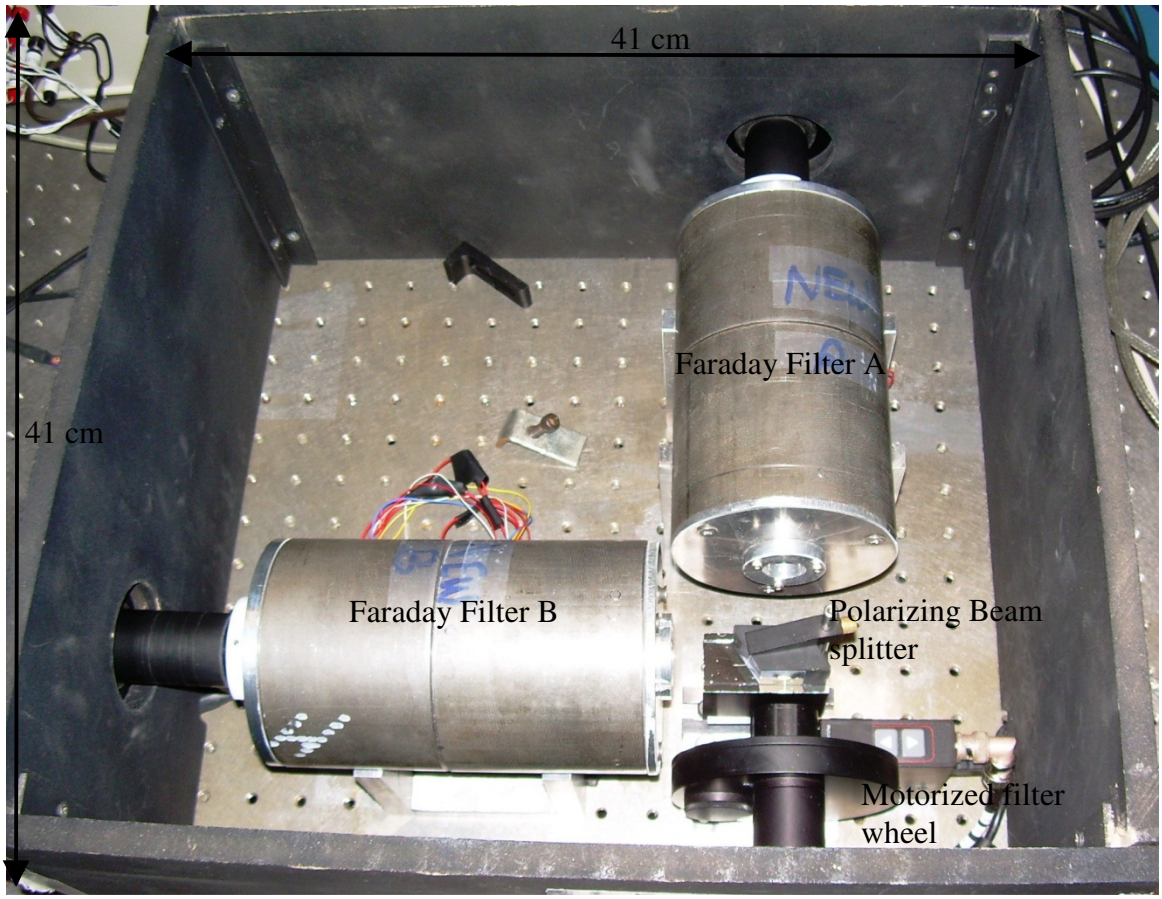

Figure 5.2: Photograph of the FFBS, with approximate dimensions and components indicated. Not shown are the two 40% quantum efficiency PMT's, which are outside the enclosure. The fiber from the telescope is also not shown; it enters the FFBS enclosure and the filter wheel from the bottom of the photograph.

specially angled side faces to produce two orthogonal linear polarization outputs at a separation of 90 degrees. This replaces the input linear polarizers typically needed in a Faraday filter. Each polarization then passes through a Faraday filter (including a Glan-Thompson linear polarizer), and then is focused by a 75 mm focal length lens onto a 40% quantum efficiency PMT (Hamamatsu model H7141).

The fiber coupling of the light from the telescope is an important aspect of determining the expected nightglow signal, so it will be discussed in detail here. Etendue is a conserved quantity in an optical system—it is defined as the product of the area of

the entrance pupil and the solid angle of the source subtended by the pupil. However, since it is a conserved quantity, this relationship holds for all the optical surfaces in the system, and the optical surface with the smallest available étendue will effectively limit the light-gathering power of the entire system. The limiting surface in the telescope receiving system for the spectrometer is the fiber used to couple the light from the telescope into the spectrometer. Currently, this fiber has a numerical aperture (NA) of 0.37, and a core diameter of 1.5 mm. For a particular NA, the solid angle subtended is $\pi(NA)^2$. Thus, the étendue for the telescope collection system is limited to $7.6\text{x}10^{-3}$ cm^2sr (The Celestron 14 has f/# = 11, while the Starsplitter telescope has f/# = 4.0). This étendue will determine the fraction of the total nightglow that will be collected by the telescope receiving system. With the nightglow intensity measured in rayleighs (chapter 2), the number of photons collected by the telescope receiving system is:

$$\frac{2\times10^{7}\ \text{photons}}{4\pi\,\text{cm}^2\ \text{s sr}}\,7.6\times10^{-3}\ \text{cm}^2\text{sr} = 1.21\times10^{4}\,\frac{\text{photons}}{\text{s}} \qquad (5.1)$$

assuming the annual minimum nightglow signal of 20 R (Chamberlain 1961).

This signal will be attenuated by the optics in the telescope receiving system and the spectrometer itself. Since ratios of measured PMT collected counts are needed, these transmission factors have not been individually determined for every optical surface since many will cancel out in the data analysis procedure. Instead, only the relative passive transmission between the two channels of the FFBS γ is determined as discussed in 5.4.

5.2 Data recording and electronic control

The FFBS can operate in one of two modes: either stand-alone or with the lidar. The details of both operation modes are given below.

5.2.1 Stand-alone operation mode

The electronic control system for the FFBS relies on a LabVIEW program running on a PC for timing and data recording. This PC contains the FAST ComTec P7882 photon-counting board (hereafter referred to as "ComTec card"). This PC is also connected to the motorized 6-position filter wheel (ThorLabs FW102B)—a USB connection triggers the filter wheel to change position, while a BNC connection from the output of the filter wheel to the ComTec card acts as the trigger signal for the card.

Timing of the data collection is controlled by LabVIEW. At the start of data collection, communication is established between LabVIEW and both the filter wheel and ComTec card. The ComTec card is then set to be ready to acquire data as soon as it receives a BNC trigger signal. LabVIEW then moves the filter wheel to position 1, containing the D_1 interference filter. When the filter wheel arrives at the proper position, it sends an output signal through the BNC cable to trigger the ComTec card to begin to acquire data.

The ComTec card is designed for time of flight photon counting. This is not needed for the FFBS, since there is no timing of the nightglow signal. Therefore, the card is set to have a certain number of range bins with a duration such that the total collection time is 60 seconds. This is achieved by using 2048 range bins, each with 29.29 ms duration.

When the card begins acquiring data, the LabView program allows the card to collect data for 60 s, and then the ComTec card is stopped. The LabView program moves the filter wheel to position 2 (the D_2 interference filter) and data is recorded for another 60 s in the same manner.

Since the dark counts of the PMT, which are substantial, could affect the results, and there is no inherent "background subtraction" that can happen from the nightglow measurements alone, after nightglow data is recorded from both D_1 and D_2 interference filters, a 30 second dark count measurement is made. This is achieved by a solid black disk in position 3 of the filter wheel. Then, the filter wheel is moved to position 4, for a minute-long 532 nm interference filter measurement. This is to estimate the sky background for elimination or correction of data sets contaminated by clouds or moonlight. After this background measurement, the program saves the recorded counts in each filter wheel position and from each PMT to a text file. The filter wheel is then moved to position one, and this process repeats. The parameters recorded are: measurement start time, signals A and B from the D_1 filter, signals A and B from the D_2 filter, the dark count measurement from PMTs A and B, and the 532 nm interference filter background estimation measurements for channels A and B. One data file is produced for an entire night's measurement—so each row in the program represents one, 3 minute and 30 second data set, with the nine parameters listed above recorded for each set.

5.2.2 Simultaneous lidar operation mode

Unfortunately, contamination from low-altititude lidar beams scattering (most likely from within the telescope room) was present when using the FFBS in the "stand-alone" mode while the lidar was operating simultaneously. Evidence for this assertion is supported by the results presented in figure 5.3, which shows spectrometer data, taken simultaneously with lidar data, binned in the same way as the lidar data. Significantly higher signal levels are recorded in the first several time bins, suggesting a correspondence to the low-altitude laser scattering (these increased levels are not seen when the lidar beam is not propagated into the atmosphere). Therefore, by using the lidar trigger and time bins of the same length as the lidar, the first 4 bins could be ignored in the FFBS data recording and the spectrometer data collected will be free of lidar contamination.

The instrument operates in much the same way as in section 5.2.1, the major difference being the source of the trigger signal and the number and length of the time bins used for the ComTec card. Instead of the filter wheel BNC output, this mode for simultaneous lidar operation utilizes the 50 Hz TTL ComTec count through signal for the lidar system (see Chapter 6). This requires a shorter time bin length of 1μs (the same as that used for lidar data taking), and the number of bins is increased to compensate. The filter wheel changes position while the Optech and ComTec lidar programs are saving data and the lidar LabView laser locking program performs a 1 s Doppler free scan of the ring laser, which is more than enough time for the spectrometer filter wheel to change positions using the USB connection. The data taking sequence is similar to the stand-alone mode: one minute for D_1, one minute for D_2, one minute for dark counts, and one

minute for the 532 nm background estimation. The reason for the one minute dark count measurement, as opposed to 30 s for the stand alone mode, is to keep the FFBS and lidar synchronized.

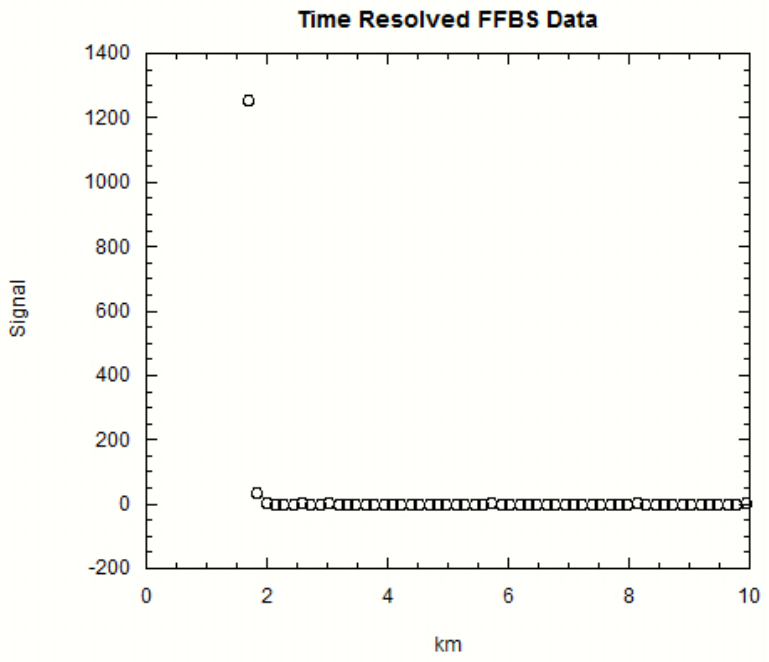

Figure 5.3: Time-resolved FFBS data collected while the lidar was running. Each point represents one range bin 150 m in length, or 1 μs. As can be seen, only the first 2-3 time bins have contamination from the lidar signal. Note that the first range bin is around 1.6 km; this is because Fort Collins is 1570 m altitude.

5.3 Data analysis

5.3.1 Reduction of FFBS signals

The data file produced by LabVIEW has already been described in section 5.2. Each individual, recorded integrated measurement is denoted as *s*. To increase the signal to noise ratio, these measurements are averaged over a longer period of time (for example, 30 minutes). Since some nights have different integration times for sky measurements and dark counts, therefore the dark count signals are corrected for the ratio

of those times. This yields six measured, averaged, dark count subtracted rates, denoted *S*:

$$S_j = \frac{\sum_{i=1}^{N} (s_j)_i}{N} - \left(\frac{t_s}{t_d N} \sum_{i=1}^{N} (s_d)_i \right) \tag{5.2}$$

where *j*= A1, B1, A2, B2, A532, or B532.

The *S* values are the nightglow intensities reduced by the fractional transmission of the spectra through the interference and Faraday filters and the other passive optical losses in the atmosphere and within the system. The fractional transmission, or *c*-coefficient, is the frequency integral of the product of the interference filter transmission, Faraday filter transmission, and the theoretical nightglow spectra described in chapter 3; it can be expressed as a coefficient *c* and is determined numerically. The interference and Faraday filter measurement scheme is described in section 5.4.

While knowledge of transmission factors from the nightglow layer to the PMT would be required for absolute nightglow intensity measurements, those measurements are difficult to make and are unnecessary if only the ratio of two nightglow signals is of interest. In practice, therefore, the ratio passive transmission of one Faraday filter beam path to the other can be utilized, and one nightglow signal can be simply written in terms of the other. Any common transmission, such as that through the atmosphere, will be the same for all signals and will cancel out in the ratio. In my analysis, I will calculate the ratio of Faraday filter A to Faraday filter B. Because of this, all of the nightglow intensities determined from FFBS data are relative, and differ from the absolute intensity in rayleighs by a multiplicative constant.

The four nightglow signals S can be written in terms of the nightglow intensities I and the laboratory-determined c-coefficients (defined earlier) and γ value:

$$\begin{aligned}
\frac{S_{A1}}{\gamma} &= I_{1A}c_{A1A} + I_{1X}c_{A1X} \\
\frac{S_{A2}}{\gamma} &= I_{2A}c_{A2A} + I_{2X}c_{A2X} \\
S_{B1} &= I_{1A}c_{B1A} + I_{1X}c_{B1X} \\
S_{B2} &= I_{2A}c_{B2A} + I_{2X}c_{B2X}
\end{aligned} \tag{5.3}$$

In the subscripts, the letter A or B represents the Faraday filter used in that measurement, the numeral 1 or 2 represents whether the D_1 or D_2 interference filter was in place while the measurement was made, and X (ground) and A (excited) represent the state of NaO used in the nightglow photon production. From these four measured, dark count subtracted signals, the values of the four nightglow intensities I can be calculated:

$$\begin{aligned}
I_{2X} &= \frac{c_{A2A}S_{B2} - \frac{c_{B2A}}{\gamma}S_{A2}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}} \\
I_{2A} &= \frac{\frac{c_{B2X}}{\gamma}S_{A2} - c_{A2X}S_{B2}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}} \\
I_{1X} &= \frac{\frac{c_{B1A}}{\gamma}S_{A1} - c_{A1A}S_{B1}}{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}} \\
I_{1A} &= \frac{c_{A1X}S_{B1} - \frac{c_{B1X}}{\gamma}S_{A1}}{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}
\end{aligned}, \tag{5.4}$$

From these intensities, the following quantities can be determined:

$$R_{DX} = \frac{I_{2X}}{I_{1X}} = \frac{c_{A2A}S_{B2} - \frac{c_{B2A}}{\gamma}S_{A2}}{\frac{c_{B1A}}{\gamma}S_{A1} - c_{A1A}S_{B1}} \frac{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}}$$

$$R_{DA} = \frac{I_{2A}}{I_{1A}} = \frac{\frac{c_{B2X}}{\gamma}S_{A2} - c_{A2X}S_{B2}}{c_{A1X}S_{B1} - \frac{c_{B1X}}{\gamma}S_{A1}} \frac{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}} \tag{5.5}$$

$$R_{D2} = \frac{I_{2A}}{I_{2X}} = \frac{\frac{c_{B2X}}{\gamma}S_{A2} - c_{A2X}S_{B2}}{c_{A2A}S_{B2} - c_{B2A}\frac{S_{A2}}{\gamma}}$$

$$R_{D1} = \frac{I_{1A}}{I_{1X}} = \frac{c_{A1X}S_{B1} - \frac{c_{B1X}}{\gamma}S_{A1}}{c_{B1A}\frac{S_{A1}}{\gamma} - c_{A1A}S_{B1}} \tag{5.6}$$

$$I_2 = I_{2X} + I_{2A} = \frac{(c_{A2A} - c_{A2X})S_{B2} + \frac{(c_{B2X} - c_{B2A})}{\gamma}S_{A2}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}}$$

$$I_1 = I_{1X} + I_{1A} = \frac{(c_{A1X} - c_{A1A})S_{B1} + \frac{(c_{B1A} - c_{B1X})}{\gamma}S_{A1}}{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}} \tag{5.7}$$

$$R_D = \frac{I_2}{I_1} = \frac{(c_{A2A} - c_{A2X})S_{B2} + \frac{(c_{B2X} - c_{B2A})}{\gamma}S_{A2}}{(c_{A1X} - c_{A1A})S_{B1} + \frac{(c_{B1A} - c_{B1X})}{\gamma}S_{A1}} \frac{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}} \tag{5.8}$$

The quantities in equations (5.5)-(5.8) can be related to [O]/[O_2] and other aspects of the modified Chapman Mechanism using equations found in chapter 3. In particular, R_{D1} and R_{D2} are directly proportional to [O]/[O_2], and R_{DX} and R_{DA} are, respectively, the NaO(X) and NaO(A) contributions to the total R_D.

5.3.2 Error propagation

The values in equations (5.5)-(5.8) will have error propagated through the calculation from various sources: photon counting error from nightglow received signals, and error arising in the deduction of *c*-coefficients and γ values, based on laboratory Faraday function measurements and passive loss measurements. All of these errors are statistical, and can be propagated through the calculation of intensity and ratio values.

The error in γ values is calculated by propagating the Poisson statistics through the calculation as described in section 5.5. The error in c-coefficients is a bit more complex. The error depends on the error in the Faraday filter transmission measurements. Since the transmission is then multiplied by a theoretical spectrum (hence no error) and then integrated using numerical techniques, it should be possible to estimate the error by propagating the error from the transmission measurement through the numerical integration algorithm. The algorithm used is INT_TABULATED, a procedure built-in to IDL. The numerical integration formula in INT_TABULATED is a five-point closed Newton-Cotes formula (Press, et al. 1992):

$$\sum_{i=5,10,\ldots}^{5000} \frac{2}{45}\Delta\nu\left[7\left((Tg)_{i-4}+(Tg)_{i}\right)+32\left((Tg)_{i-3}+(Tg)_{i-1}\right)+12(Tg)_{i-2}\right]$$

where *Tg* are the values of the filter transmission and nightglow spectrum product, evaluated at frequency points spaced with an interval Δν, and the sum is performed for *i*=5,10,…5000. There are 5000 points in the frequency scale (from -5 GHz to 5 GHz, with a 0.002 GHz step size). From propagation through the Newton-Cotes formula, the error in the *c* coefficients is:

$$\sqrt{\sum_{i=5,10\ldots}^{5000}\frac{2\Delta\nu}{45.0}\sqrt{49\left(\sigma_{Tg(ii-4)}{}^{2}+\sigma_{Tg(ii)}{}^{2}\right)+1024\left(\sigma_{Tg(ii-3)}{}^{2}+\sigma_{Tg(ii-1)}{}^{2}\right)+144\sigma_{Tg(ii-2)}{}^{2}}} \tag{5.9}$$

where σ is the error in z at each frequency point, calculated from the standard deviation of the mean of the average of successive Faraday filter transmission measurements (typically, more than 10 scans are averaged for each coefficient determination).

The error in the signal measurements is propagated through the calculation from the individual s measurements of equation (5.2). Since photon counting follows Poisson statistics, the error in s is $\sqrt{s}$. Therefore, the error in the averaged S signals is:

$$\Delta S_j = \frac{1}{N}\sqrt{\sum_{i=1}^{N}(s_j)_i + \left(\frac{t_s}{t_d}\right)^2 \sum_{i=1}^{N}(s_d)_i} \tag{5.10}$$

where j = A1, B1, A2, B2, A532, or B532, and s are the individual minute (or 30 second) measurements. The subscript d represents the dark count measurement.

Since each of the sources of error is independent, it can be propagated through equation (5.3) separately, and then added in quadrature to get a total error value. This process is useful because it can give a sense of the dominant source of error in the experiment and suggestions for future improvement. Since the equations for error bar assessment quickly become quite complex, only representative terms were calculated and the equations will not be detailed here. A set of data (UT day 10076; to be presented in chapter 7) was utilized to determine the dominant sources of error for a representative term, R_{DX}. Error from photon noise dominates; in fact, the lower signal levels contribute larger share of the error. The resulting error from the c coefficients and γ factor is, at most, 2% of the largest error due to the nightglow signals; this leads to a negligible increase in the total error over that of photon noise alone. When the error of two of the c coefficients was propagated through the calculation of R_D, it was again found to have a negligible contribution to the total error. When accounting for possible Faraday filter

transmission function variation, up to ±0.05 is factored into the error of the *c* coefficients, the error introduced by the *c* coefficients is still negligible.

Since photon noise of the four measured signals is the dominant source of error, I will only show the equations containing the propagation of photon noise from equation (5.10). Propagating this error through equation (5.4) gives:

$$\begin{aligned}
\Delta I_{2X} &= \sqrt{\left(\frac{1}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}}\right)^2\left[{c_{A2A}}^2(\Delta S_{B2})^2 + \frac{{c_{B2A}}^2}{\gamma^2}(\Delta S_{A2})^2\right]} \\
\Delta I_{1X} &= \sqrt{\left(\frac{1}{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}\right)^2\left[{c_{A1A}}^2(\Delta S_{B1})^2 + \frac{{c_{B1A}}^2}{\gamma^2}(\Delta S_{A1})^2\right]} \\
\Delta I_{2A} &= \sqrt{\left(\frac{1}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}}\right)^2\left[{c_{A2X}}^2(\Delta S_{B2})^2 + \frac{{c_{B2X}}^2}{\gamma^2}(\Delta S_{A2})^2\right]} \\
\Delta I_{1A} &= \sqrt{\left(\frac{1}{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}\right)^2\left[{c_{A1X}}^2(\Delta S_{B1})^2 + \frac{{c_{B1X}}^2}{\gamma^2}(\Delta S_{A1})^2\right]}
\end{aligned}. \qquad (5.11)$$

Error propagation through equations 5.5 through 5.8 gives:

$$\begin{aligned}
\Delta R_{DX} &= \sqrt{{R_{DX}}^2\left[\frac{{c_{A2A}}^2(\Delta S_{B2})^2 + \frac{{c_{B2A}}^2}{\gamma^2}(\Delta S_{A2})^2}{\left(c_{A2A}S_{B2} - \frac{c_{B2A}}{\gamma}S_{A2}\right)^2} + \frac{{c_{A1A}}^2(\Delta S_{B1})^2 + \frac{{c_{B1A}}^2}{\gamma^2}(\Delta S_{A1})^2}{\left(\frac{c_{B1A}}{\gamma}S_{A1} - c_{A1A}S_{B1}\right)^2}\right]} \\
\Delta R_{DA} &= \sqrt{{R_{DA}}^2\left[\frac{{c_{A2X}}^2(\Delta S_{B2})^2 + \frac{{c_{B2X}}^2}{\gamma^2}(\Delta S_{A2})^2}{\left(\frac{c_{B2X}}{\gamma}S_{A2} - c_{A2X}S_{B2}\right)^2} + \frac{{c_{A1X}}^2(\Delta S_{B1})^2 + \frac{{c_{B1X}}^2}{\gamma^2}(\Delta S_{A1})^2}{\left(c_{A1X}S_{B1} - \frac{c_{B1X}}{\gamma}S_{A1}\right)^2}\right]}
\end{aligned} \qquad (5.12)$$

$$\Delta R_{D2} = \sqrt{{R_{D2}}^2\left[\frac{\frac{{c_{B2X}}^2}{\gamma^2}(\Delta S_{A2})^2 + {c_{A2X}}^2(\Delta S_{B2})^2}{\left(\frac{c_{B2X}}{\gamma}S_{A2} - c_{A2X}S_{B2}\right)^2} + \frac{{c_{A2A}}^2(\Delta S_{B2})^2 + \frac{{c_{B2A}}^2}{\gamma^2}(\Delta S_{A2})^2}{\left(c_{A2A}S_{B2} - \frac{c_{B2A}}{\gamma}S_{A2}\right)^2}\right]}$$

$$\Delta R_{D1} = \sqrt{{R_{D1}}^2\left[\frac{{c_{A1X}}^2(\Delta S_{B1})^2 + \frac{{c_{B1X}}^2}{\gamma^2}(\Delta S_{A1})^2}{\left(c_{A1X}S_{B1} - \frac{c_{B1X}}{\gamma}S_{A1}\right)^2} + \frac{\frac{{c_{B1A}}^2}{\gamma^2}(\Delta S_{A1})^2 + {c_{A1A}}^2(\Delta S_{B1})^2}{\left(\frac{c_{B1A}}{\gamma}S_{A1} - c_{A1A}S_{B1}\right)^2}\right]} \quad (5.13)$$

$$\Delta I_2 = \sqrt{\left(\frac{1}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}}\right)^2\left[(c_{A2A} - c_{A2X})^2(\Delta S_{B2})^2 + \frac{(c_{B2X} - c_{B2A})^2}{\gamma^2}(\Delta S_{A2})^2\right]}$$

$$\Delta I_1 = \sqrt{\left(\frac{1}{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}\right)^2\left[(c_{A1X} - c_{A1A})^2(\Delta S_{B1})^2 + \frac{(c_{B1A} - c_{B1X})^2}{\gamma^2}(\Delta S_{A1})^2\right]} \quad (5.14)$$

$$\Delta R_D = \sqrt{{R_D}^2\left[\frac{(c_{A2A} - c_{A2X})^2(\Delta S_{B2})^2 + \frac{(c_{B2X} - c_{B2A})^2}{\gamma^2}(\Delta S_{A2})^2}{\left[(c_{A2A} - c_{A2X})S_{B2} + \frac{c_{B2X} - c_{B2A}}{\gamma}S_{A2}\right]^2} + \frac{(c_{A1X} - c_{A1A})^2(\Delta S_{B1})^2 + \frac{(c_{B1A} - c_{B1X})^2}{\gamma^2}(\Delta S_{A1})^2}{\left[\frac{c_{B1A} - c_{B1X}}{\gamma}S_{A1} + (c_{A1X} - c_{A1A})S_{B1}\right]^2}\right]} \quad (5.15)$$

These formulas will be used for error bar calculation in Chapter 7.

5.4 Faraday filter design and temperature control

The Faraday filter design for the FFBS is shown in figure 5.4. The main components are the steel tube which holds all of the components, the 12 magnet bars (6 at each end), which form circles at each end of the steel tube. The north pole is the outside diameter of one ring, while the south pole is the outside diameter of the other ring; this forms the uniform magnetic field of 2500 G.

The temperature is controlled, as mentioned in chapter 4, in two regions: the main body of the cell, and the tip containing the Na reservoir. The main body of the cell is heated by four cartridge heaters, two in either end of the aluminum oven encircling the cell, inserted axially in the oven. The temperature is measured by a cylindrical RTD placed axially in one end of the oven. The temperature is controlled by a proportional-integral-differential (PID) controller manufactured by Fuji Electric. The tip region is heated by a thin film (Kapton) heater wrapped around the aluminum "cup" oven which surrounds the tip; the temperature is measured by a thin-film RTD also is placed externally to the aluminum oven—both are held in place by a Teflon cup which acts as an insulator. The temperature is controlled by another PID temperature controller. Since the temperature for both the cell and tip are measured outside of the cell, the temperatures set at the PID controllers are not the theoretically-derived temperatures. These set temperatures are chosen to best match the experimental transmission function with the theoretical function.

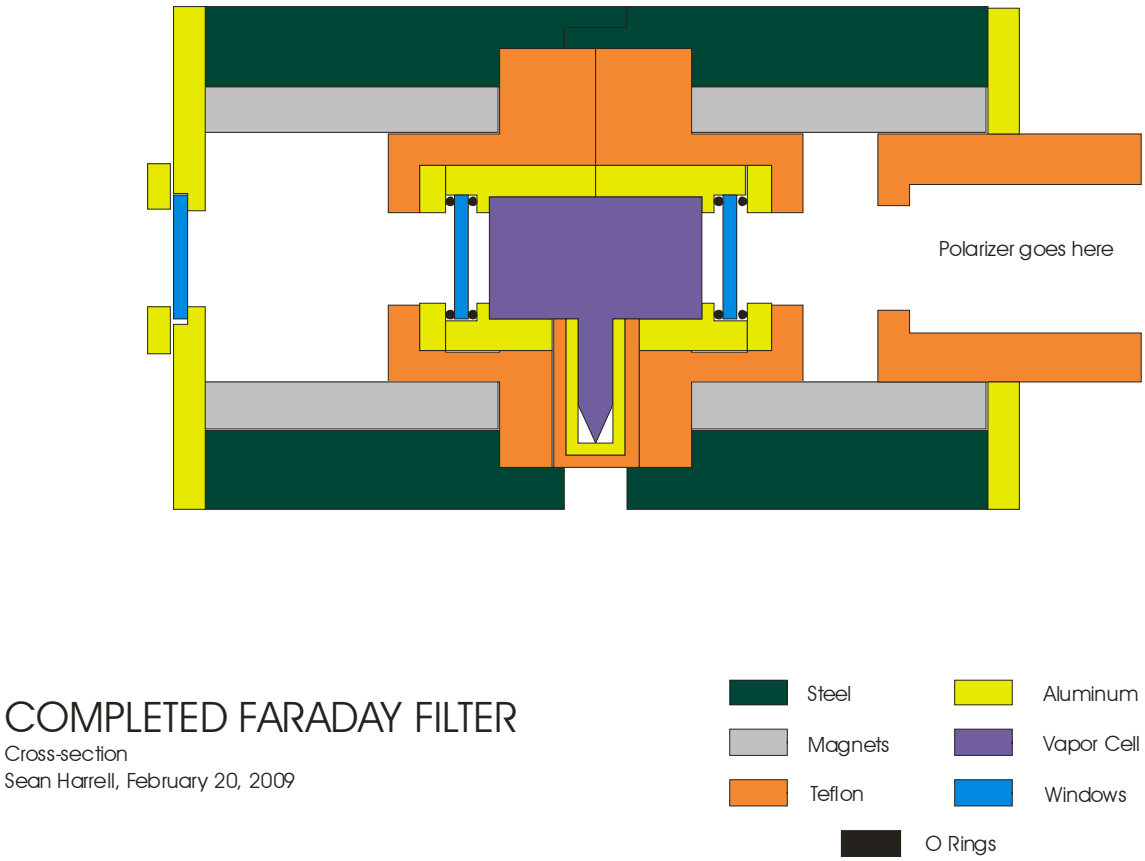

Figure 5.4: Cross-section of the Faraday filter design. The light enters the Faraday filter from the left, and exits from the right, where "Polarizer goes here" indicates the location of the Glan-Thompson (second) polarizer.

5.5 Scanning the Faraday and Interference filters

For accurate measurements with the FFBS, knowledge of the Faraday filter and interference filter transmission functions on a high-resolution frequency scale is required. Measurements are made using various parts of the lidar transmitter system which will be described in detail in Chapter 6.

The scanning system is shown in figure 5.5. Narrowband laser light is produced by the ring dye laser. The light is then monitored for frequency scale and power fluctuations, and a portion of the light enters the Faraday filter or interference filter to make the transmission function measurement.

The Doppler free spectroscopy sub-system is described in chapter 6.3.1. It produces a sodium spectrum with several sharp features at fixed frequency points. The Fabry-Perot interferometer has a 750 MHz free spectral range; our single-mode ring dye laser thus produces a peak at every 750 MHz as it is scanned in frequency; however these peaks are not at any particular absolute frequency. An example of each photodiode or PMT recorded scan signal is shown in figure 5.6.

The data is recorded with a LabView program through a BNC-2090 Board. Twenty-four thousand points are recorded over 2.0 minutes using a 50 Hz TTL signal to trigger the 2090 Board. The ring dye laser is set to perform a single, internally-controlled scan over a range of at least 16 GHz. The LabView program and the laser scan are started simultaneously, and the LabView program records all signal levels into a text file.

A program written in the IDL language is used to analyze the scan data. The first step of this program is to smooth the scan data to eliminate large spikes of noise, especially in the Doppler free spectrum, where it could cause problems with identification of the D_{2a} Lamb dip and crossover peak. Next a linear interpolation of the Fabry Perot interferometer signal is used to determine the relative frequency scale of the data points between each peak based on the FSR of 750 MHz. This scale is then set to absolute frequency using the position of the D_{2a} or D_{1a} Lamb dip. The difference between the D_{1a} or D_{2a} Lamb dip and the crossover peak is calculated and the value is checked by the user against the accepted value; if the calculated value differs by more than ±0.001 GHz, then the scan is discarded. Otherwise, the value of the Faraday or interference filter is divided by the power normalization channel to correct for laser

power fluctuations. This gives an absolute frequency scaled, power fluctuation corrected Faraday filter scan relative transmission.

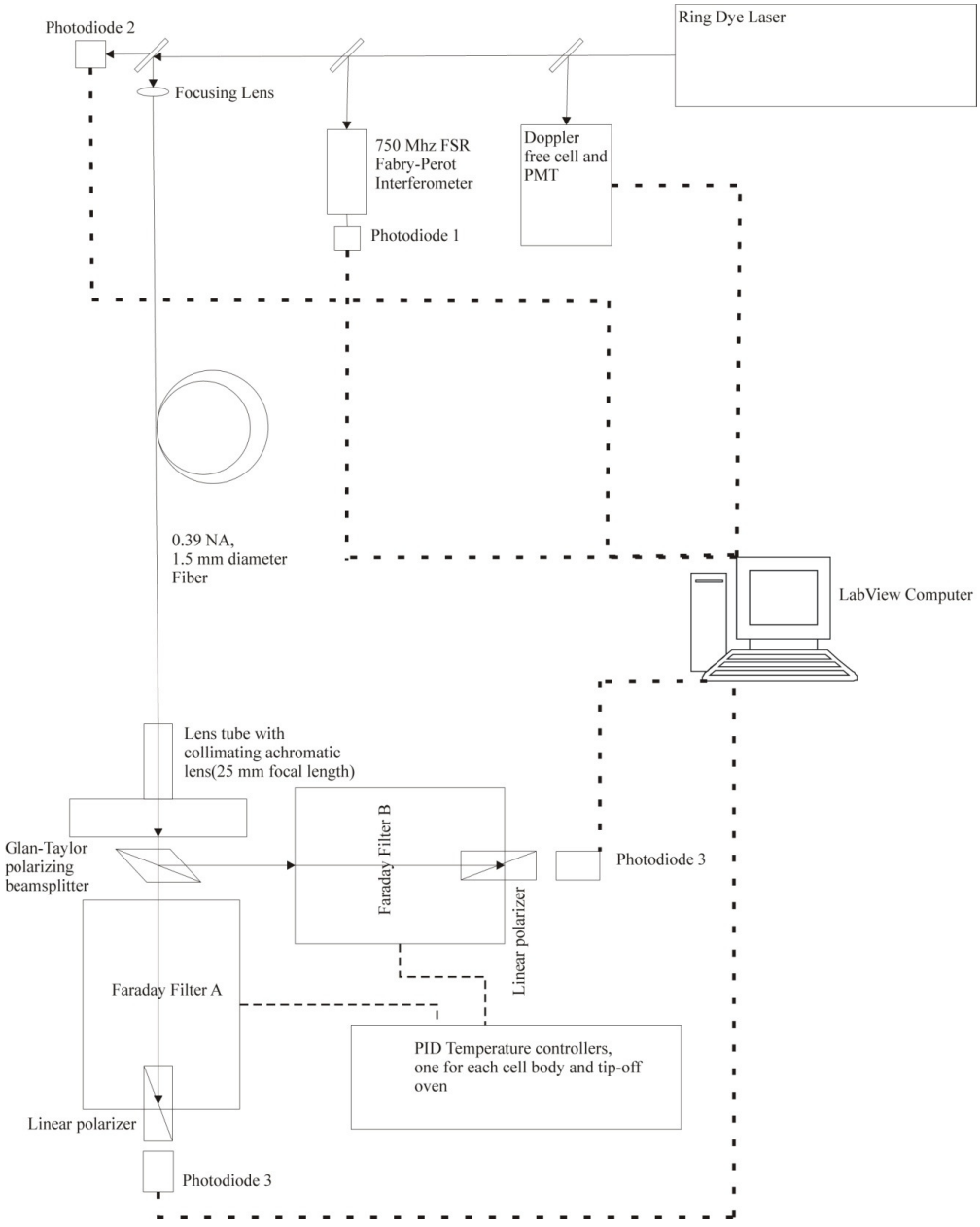

Figure 5.5: Diagram for scanning of Faraday filters (not to scale). For interference filter scanning, photodiode 3 is placed directly after the filter wheel. Absolute frequency is monitored by the Doppler free spectroscopy unit, and relative frequency is monitored by the Fabry-Perot interferometer. Laser power is monitored by photodiode 2, while the transmission function is measured by photodiode 3.

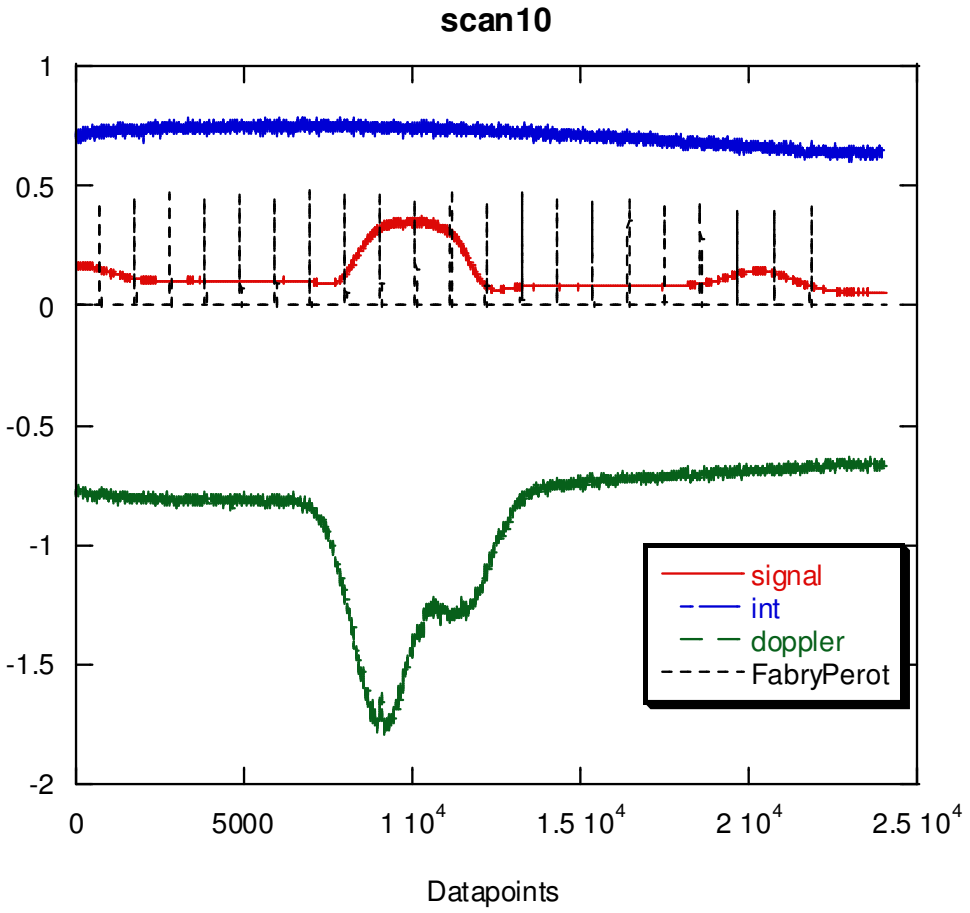

Figure 5.6: Example of a Faraday filter scan file. The black Fabry-Perot curve has peaks at every 750 MHz, this can be scaled to absolute frequency using the Doppler free spectrum (in green), which has the D_{2a} Lamb dip near the minimum voltage. The blue curve is the laser normalization channel, which is used to correct for varying laser power over the course of a scan.

The next step is to take the average of several scans to reduce noise in the scan voltage. Then, the process described in section 4.3, finding the constant point of 25% transmission, is used to figure out the absolute transmission level. The result is plotted in figure 5.7.

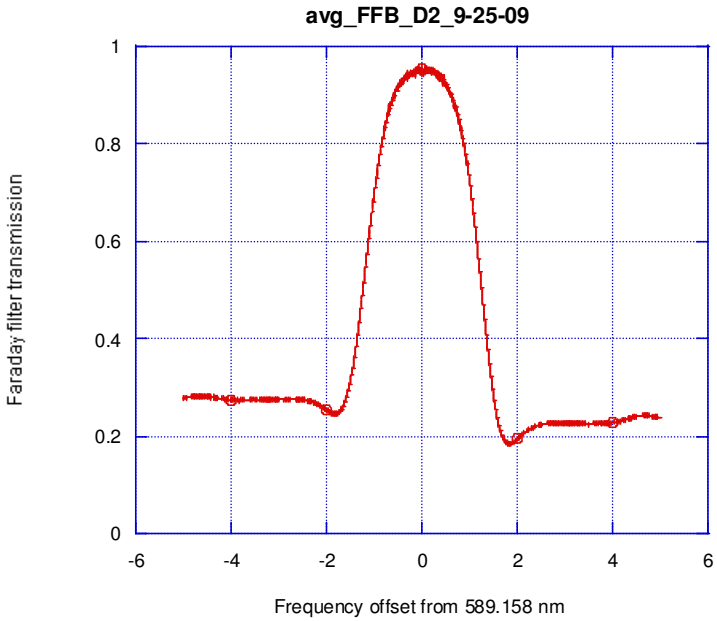

Figure 5.7: A sample analyzed average of several Faraday filter scans, corrected for absolute transmission. The frequency scale is in GHz.

Error is calculated in these scans by taking the standard deviation of the mean of the averaged values at each frequency point. The error for Faraday and interference filter scans can then be incorporated into the *c* coefficients as described in 5.3.

5.6 Measuring the γ passive transmission factor

The passive transmission factor γ is measured using the ring laser output as input into the fiber of the spectrometer as set up in figure 5.1. Using neutral density filters the power of the laser light entering the spectrometer is greatly reduced to a level acceptable for proper operation of the 40% quantum efficiency of the PMT's, which have a maximum of $1x10^6$ counts per second for linear operation. The laser is locked to either the D_{2a} or D_{1a} peak using the Doppler free spectroscopy system; typically the D_{1a} peak is used because of extremely low transmission at the D_{2a} peak for one of the Faraday filters. Once data with the laser on is collected, the laser output is blocked to get a background

measurement. The dark count and background subtracted data is then corrected for the Faraday filter transmission at the frequency at which the laser was locked. Then, the number of counts measured in PMT A can be divided by the number of counts measured in PMT B to yield the value of γ needed in equation 5.3. Since this measurement is made with the 40% photon-counting PMT's, the error is calculated from Poisson statistics (propagated through the calculation).

5.7 Determining Faraday Filter temperature settings

The FFBS can be optimized to minimize measurement error in any of the quantities in equations 5.3-5.7. This process is accomplished by changing the temperature settings (for the predetermined and fixed cell length and magnetic field) of the two Faraday filters to vary filter transmission, which in turn varies the values of the fraction of the nightglow spectra which will pass through the Faraday filter. By using the built-in IDL minimization procedure AMOEBA, the temperature settings which yield minimum error can be found. Then, the experimental temperature settings can be adjusted until the actual Faraday filter transmission functions most closely match those of the theoretical calculation. This minimization is theoretically independent of the actual amount of nightglow collected; while the absolute error value will decrease with increased nightglow signal, the temperature setting producing relative minimum error (assuming the same Faraday filter vapor length and magnetic field) can be determined. In practice, it might be more desirable to increase the Faraday filter transmission to increase the signal to noise ratio to account for undesirably low signal levels.

Unfortunately, this minimization procedure was carried out by making some incorrect theoretical assumptions regarding the chemical processes in chapter 3.

However, the temperature settings acquired still lead to reasonably low error in R_D with the correct chemical theory. After performing the minimization according to earlier theory, I get the following theoretical temperature settings for the vapor and sodium reservoir temperatures: 459 K and 454 K for filter A vapor and reservoir, respectively; and 450 K and 445 K for filter B vapor and reservoir, respectively. These theoretical Faraday transmission functions are plotted in figure 5.8, along with a corresponding set of experimental measurements. Because the actual temperatures of the ovens for the cell and Na reservoir will be different than these set values since the temperature sensor is not inside the cell, there will be some discrepancies between the theoretically predicted and experimentally measured Faraday functions. The measured Faraday functions are found to be reproducible over approximately four weeks.

It is important that this minimization process is performed. As will be shown for specific data sets and values of c coefficients, it is possible to introduce unnecessary error into the calculation with an inauspicious choice of Faraday filter temperature settings. This is the case even if this change increases the overall signal levels.

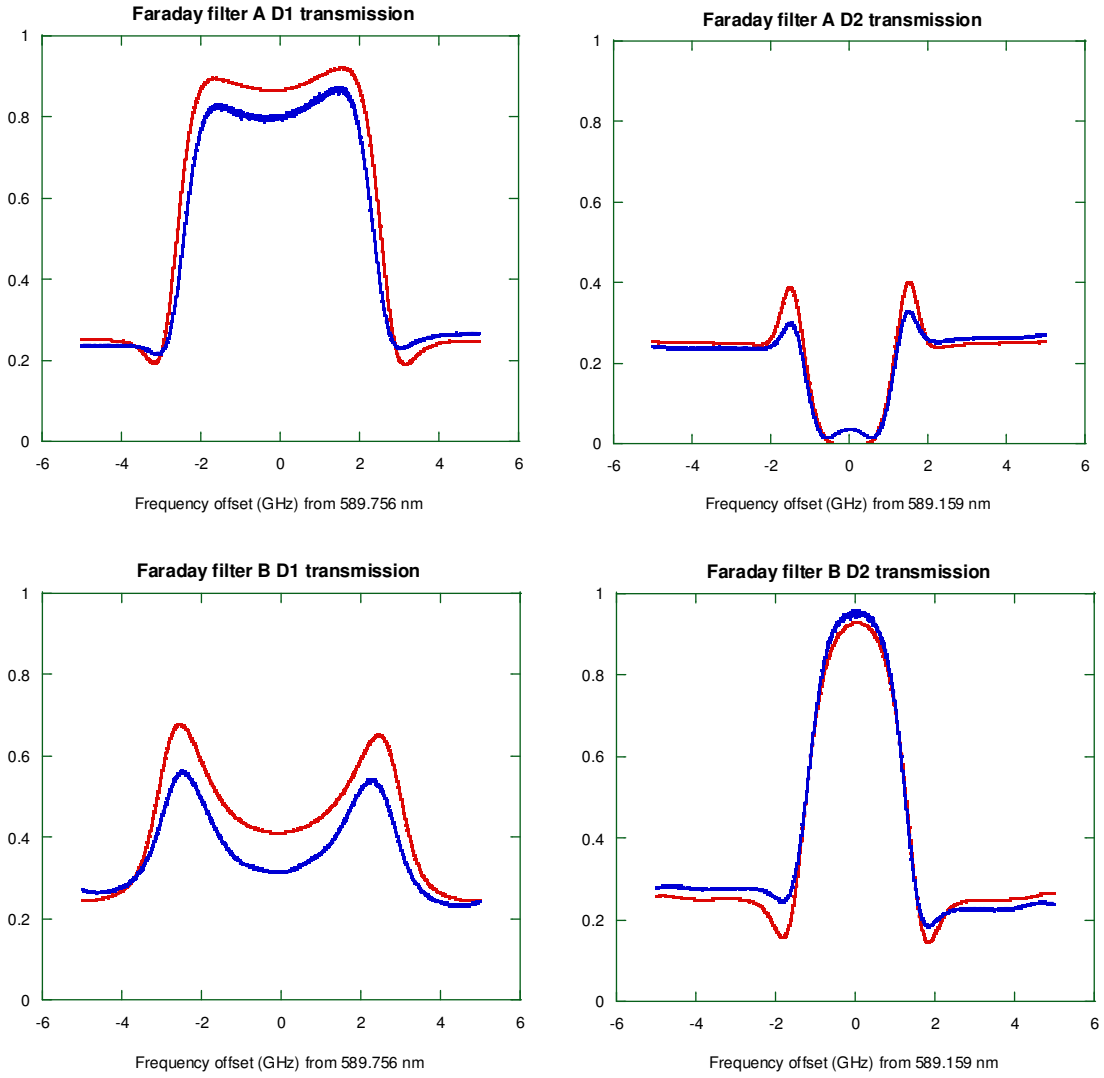

Figure 5.8: The theoretical (red) and experimental (blue) Faraday filter transmission vs frequency offset from line center for the two Faraday filters. Top Left: Faraday filter A D_1, Top Right: Faraday filter A D_2, Lower Left: Faraday filter B D_1 Lower Right: Faraday filter B D_2. The experimental data are the average of several scans.

CHAPTER 6

COLORADO STATE UNIVERSITY SODIUM LIDAR SYSTEM

While the Faraday Filter-Based Spectrometer can be operated as a stand-alone system, there are several advantages gained by operating this system alongside the CSU Sodium lidar system. These are the ability to accurately determine the Na layer height and height distribution based on lidar-determined sodium density, determine the effect of solar tides, and observe the evidence of gravity waves based on lidar temperature and wind measurements. All of these could affect the $[O]/[O_2]$ ratio, so these collocated measurements will be a part of this study. Thus, an understanding of the principles of lidar temperature, wind, and Na density measurements are important.

6.1 Introduction to the CSU Sodium Lidar System

The Colorado State University (CSU) sodium lidar system has been developed over the past twenty years to become the most advanced system of its kind in the world. Applications of technology and innovations include Doppler-free spectroscopy for accurate frequency control (She and Yu 1995), acoustooptic modulation (AOM) for accurate frequency shifting (She and Yu 1994), Faraday filters for 24-hour continuous observation (Chen, et al. 1996), and frequency "chirp" measurement to correct wind measurements (Yuan, et al. 2009). These innovations allow for determination of mesopause region temperature, winds, tides, and wintertime zonal momentum flux.

In this chapter, I will describe the CSU Sodium lidar system. This will begin with a description of the measurement technique and unique solutions employed. Then, I will describe the data analysis process. As the lidar measurement is not the main focus of this dissertation, and this lidar system has been well documented in the literature, the sections in this chapter will describe aspects of the lidar most important to the FFBS, while other aspects of the lidar can be found in Appendix B.

6.2 Lidar Basics

Lidar is similar to radar; however it operates in the visible, infrared, or ultraviolet portion of the electromagnetic spectrum. A light pulse is sent out from a source, scatters off a target, and is received by a detector. The nature of the interaction between the light and scatterer dictates what parameters may be determined; all types of pulsed lidar allow for range measurements since the time between when the light is emitted and when it is detected can easily be measured. For sodium lidar, the hyperfine spectrum of the D_2 line is probed with very narrowband laser pulses at three predetermined frequencies to yield sodium density and neutral temperature and wind in the mesopause region, since the returned signal at the three frequencies will depend on the Doppler shift due to the wind and Doppler broadening due to the temperature. The specific details of the measurement theory can be found in Appendix B.

6.3 The lidar transmitter

The lidar transmitter is sketched in figure 6.1. The heart of the lidar transmitter is a narrowband (< 1MHz), continuous-wave tunable ring dye laser (Coherent 899). This

laser is capable of producing around 400 mW of light at the D_{2a} frequencies. Due to the sensitive nature of temperature and wind measurements, the laser frequency must be actively monitored and locked to the D_{2a} peak. There must also be a way to shift the laser light to ν_+ and ν_-. It is necessary to create pulsed light of a high power to achieve a reasonable signal-to-noise ratio, and a system is needed to monitor the frequency of these transmitted pulses. These are the main sub-systems of the optical system of the CSU lidar transmitter. Since the Doppler free subsystem is critical for determining Faraday filter transmission functions (see Section 5.4), it will be described here, while the other subsystems are described in Appendix B.

6.3.1 Doppler free spectroscopy for laser locking

To accurately lock the ring laser to the correct frequency, Doppler free fluorescence saturation spectroscopy is utilized. Fluorescence saturation spectroscopy occurs when the ground state of the atoms are depleted by the incident laser. To understand this system, one can first imagine a single laser beam propagating through a heated sodium vapor cell. If the atoms are stationary, each atom would fluoresce at the transition frequency ν_0; however, the atoms at a given temperature may be moving according to the Maxwell-Boltzmann temperature distribution of velocities. Because of this distribution, some atoms will be stationary, while others will be moving with various speeds along the direction of laser propagation. For laser light at ν_0 the only atoms that will interact are those that are stationary. Atoms at a speed v can only interact with a laser of frequency ν (offset from the transition frequency) which satisfies a Doppler shift to the transition frequency ν_0.

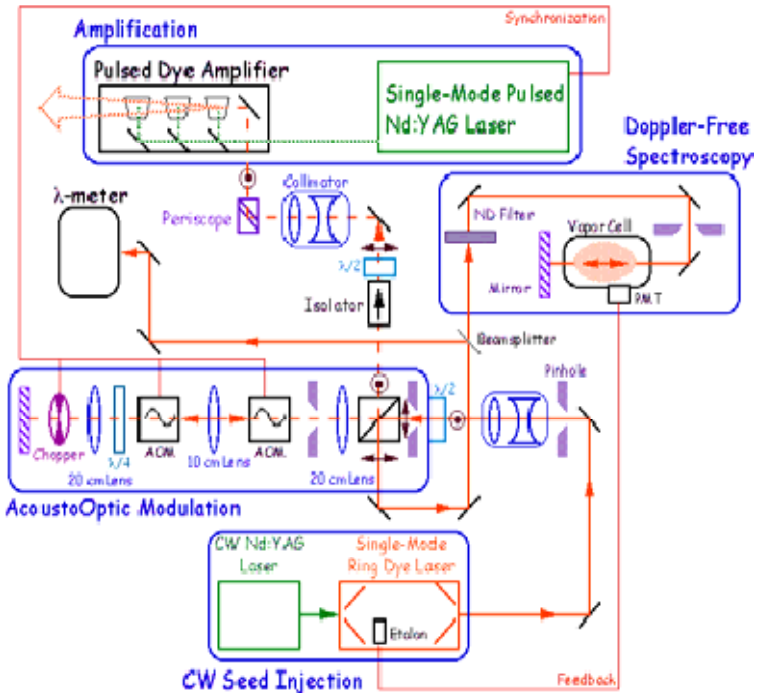

Figure 6.1: The CSU Na lidar transmitter system. Figure from Li (2005).

One can imagine placing a mirror at the output end of the cell, as shown in figure 6.1. The light beam at a frequency less than ν_0 propagating to the left (right) can only interact with a group of atoms moving to the right (left). For a laser frequency greater than ν_0, it can only interact with atoms moving in the same direction, which see a downward Doppler shift. On resonance, however, it is still only the stationary atoms that fluoresce with both counter-propagating beams; therefore, there will be saturation at on- and near-resonance, which will lead to a reduction, known as a Lamb dip, at ν_0, as shown in figure 6.2. This Lamb dip creates a sharp, repeatable peak which can easily be detected and used for laser frequency control.

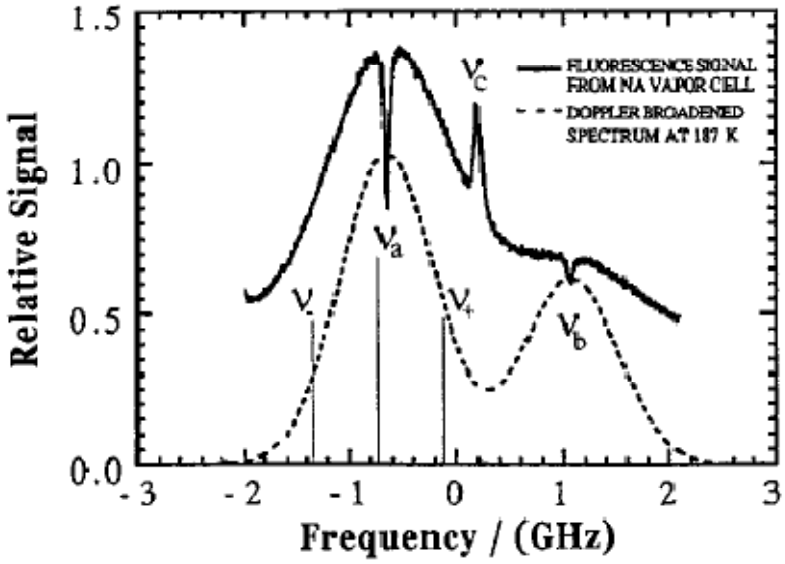

Figure 6.2: Comparison of the Doppler Free spectrum (solid) and the Doppler broadened spectrum (dashed). Note the lamb dips at the D_{2a} and D_{2b} peaks, and the crossover peak halfway in between. For details see text. From Arnold and She (2003).

This simplified picture is still incomplete; the sodium D_2 spectrum is made up of two closely spaced groups of transitions; D_{2a} and D_{2b}. These two transitions share an excited state, but utilize different, but closely spaced, ground states. At both of these peaks, there will be a lamb dip, as shown in figure 6.2. At other frequencies, if the atom is at the correct velocity and ground state to interact with a laser beam, then it will fluoresce into either of the ground states. If the atom returns to the ground state from which it started, then it can continue to interact with the laser beam. However, if it returns to the other ground state, it typically cannot—the Doppler shifted frequency of the light no longer matches the transition frequency for this ground state. This will lead to a reduction of the laser induced fluorescence (LIF) spectrum at off resonance frequencies. However, at the crossover frequency, the frequency exactly between the two resonance peaks, an atom in one ground state will interact with one laser beam. If it then relaxes into the other ground state, it will be able to interact with the other laser beam

propagating in the opposite direction. This leads to the enhanced "crossover peak" designated in the figure.

Even this three-level picture is simplified—it ignores the hyperfine structure of the excited state of the sodium atom completely. She and Yu (1995) describe this in great detail. The hyperfine structure causes the number of states and allowed transitions to increase. However, for D_2, these many transitions will be extremely closely spaced and lead to an overlapping of all the individual lamb dip and crossover resonances to the single peaks observed in figure 6.2. For the D_1 transition, the excited state ($^2P_{1/2}$) has no electric quadrupole interaction and a larger magnetic dipole interaction (B_J=0 and a larger A_J value, see Table 4.3). Therefore, there are fewer but better resolved hyperfine transitions. The D_1 Doppler free spectrum has three well resolved Lamb dips at D_{1a} and D_{1b}, and three corresponding crossover peaks (She and Yu 1995), as shown in figure 6.3. While unimportant for the operation of the lidar, this is important for the scanning and characterization of Faraday filter D_1 transmission spectra as discussed in Chapters 4 and 5. In the D_1 case, the highest of the three peaks is used at each location (D_{1a}, crossover, and D_{1b}) to determine the scan's frequency scale.

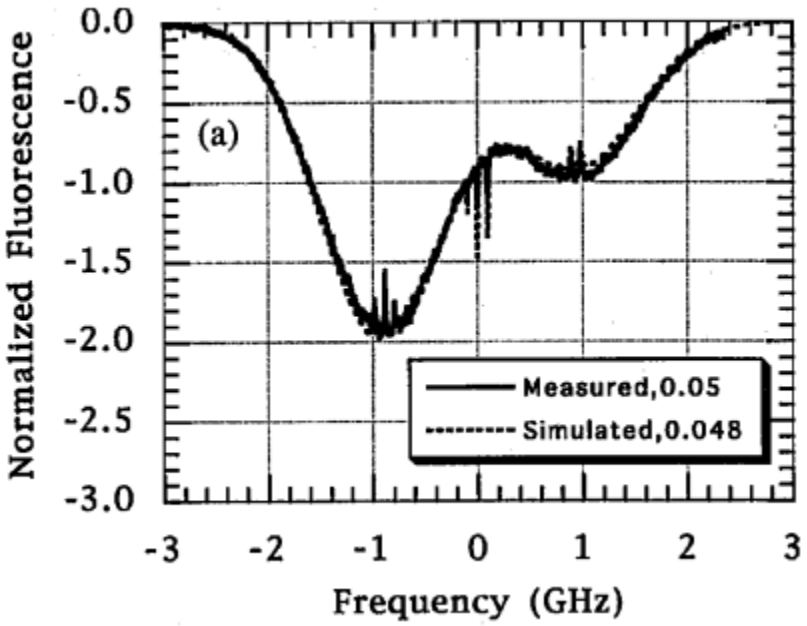

Figure 6.3: D_1 Doppler free spectrum, showing the three Lamb dips at D_{1a} and D_{1b} and the three crossover peaks. From She and Yu (1995).

In practice, the ring dye laser is manually scanned to the D_{2a} lamb dip, and then the laser is locked to this position using a PC-based lock-in amplifier feedback mechanism. This holds the laser at that frequency while the lidar returns are collected, then the PC automatically scans the laser, and relocks it to this lamb dip position. Thus, the laser is always at the correct frequency.

6.3.2 Transmitter conclusions

After the light exits the pulsed dye amplifier in Figure 6.1, it is steered upward into the sky using high-reflection dielectric mirrors. The pointing of the mirrors may be adjusted, and is from time to time. The standard wind and temperature measurement scheme uses two beams pointed 30° and 20° off zenith in the north and east directions, respectively. This allows for temperature, Na density, and zonal and meridional

horizontal wind measurements. In the winter, we point a third beam 20° off zenith to the west (along with the 20° east beam) to measure zonal momentum flux as described by Acott (2009). Dielectric plate beam splitters are used to divide the power into two or three based on the requirements of the receiver and/or the experiment. The pulses of light are transmitted to the sodium layer, interact with the Na atoms, and return to be collected and analyzed with the lidar receiver system, which is described in Appendix B.

6.4 The electronic control system

While all of the components discussed thus far are necessary, they are useless without the heart of the lidar system, the electronic control system. The electronics system produces and processes electronic signals to "time" all of the various components to work in harmony. These electronic components are also necessary to operate the FFBS, so they will be detailed here.

Figure 6.4 shows the timing diagram of the signals. The master clock of the lidar is the mechanical chopper in the 40% QE PMT system for detecting the lidar returns discussed Appendix B. This 2-slot chopper blade operates at a frequency of 100Hz (which reduces to a 50 Hz chopping rate for each of the two PMT's). This 100 Hz TTL signal enters the electronic circuit box (Acott 2009), where it is divided to a 50 Hz signal, which triggers the pulsed Nd:YAG laser of the lidar transmitter. AOM 1, AOM 2, and AOM chopper control are generated by the logic circuit from the 100 Hz signal chopper signal, and are used to turn the Acoustooptic modulator (AOM) crystals on and off at the proper times (50 Hz) to allow the passage of laser pulses of the three laser frequencies in the proper sequence. AOM chopper control is used to control the AOM chopper

synchronizer, which keeps the AOM chopper synchronized with the Nd:YAG laser output. The AOM subsystem is described in Appendix B.

Labview and Optech lowhold signals are generated by the Labview and Optech control and data taking computers, respectively. These signals ensure that the data taking does not begin to occur until the previous Optech lidar data file is saved. Labview lowhold does not allow data taking to start until the interfile scan of the Dopper free spectrum is complete and the laser is relocked to the D_{2a} peak.

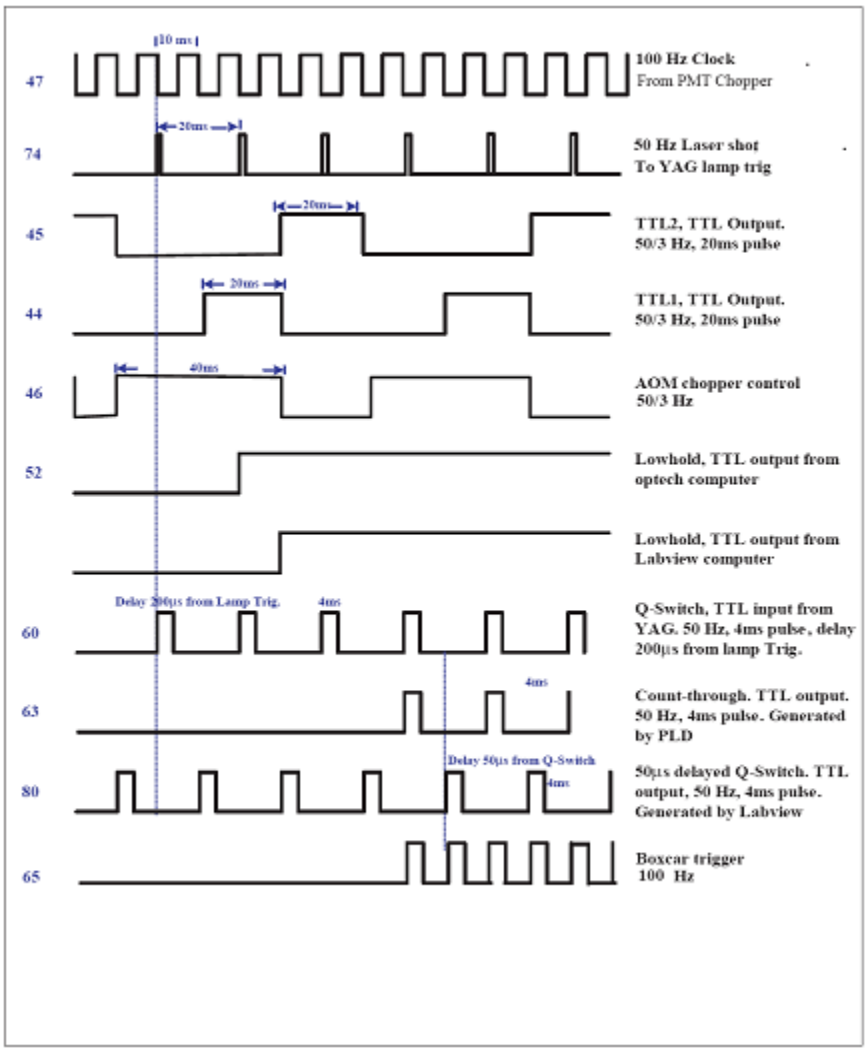

Figure 6.4: The electronics timing diagram for the CSU lidar system. For details, see the text. "Count-through" is split to trigger both the lidar and FFBS when the FFBS is run simultaneously with the lidar.

Once both lowhold signals are high, the logic circuit will allow count-through to be generated. This signal is a doubled version of the Q-switch signal, which is generated by the Nd:YAG laser when each pulse is generated. The count through signal is used for the Optech card and the Comtec photon counting cards. The count through signal is split and used as the FFBS Comtec card trigger when the FFBS and lidar operate simultaneously.

Q-switch is also used to generate several other signals: the 50 µs delayed Q-switch, and the boxcar trigger. The 50 µs delayed Q-switch is used to trigger the blanking circuit of the 20% quantum efficiency PMT's used in the lidar receiver. The boxcar trigger is a doubled version of Q-switch. It is used to trigger the chirp signal collection boxcars/gated integrators. The chirp boxcars sample the I_2 and reference photodiodes twice per laser pulse; one during the pulse transmission, and then again after the pulse for a background (and electronic bias) subtraction measurement. More details on the chirp subsystem are in Appendix B.

6.6 Lidar results and the FFBS

The lidar provides data that can complement the measurements made by the FFBS. The most important of these is the sodium density results. Figure 6.5 shows a sample contour plot of sodium density. This data showing the sodium density vs. time and altitude could be helpful in determining the altitude dependence of the FFBS nightglow measurements. This altitude dependence could show how the nightglow parameters depend upon the height and thickness of the layer, and thus, how $[O]/[O_2]$ depends on altitude as well. This information cannot be determined by the FFBS alone,

so the co-location with the CSU lidar increases the ability of the FFBS to make novel and useful measurements. Unfortunately, Na density measurements alone are not enough to determine the altitude variations of the nightglow; knowledge of the variable O_3 profiles is also necessary to completely describe the nightglow altitude dependence.

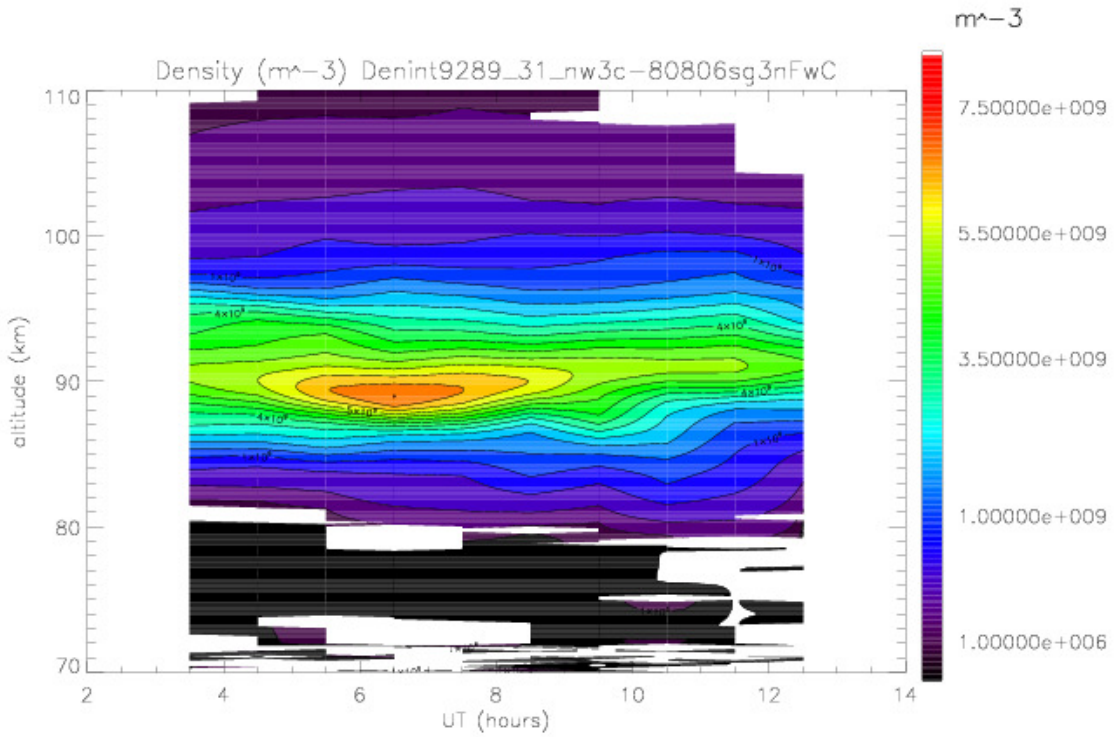

Figure 6.5: CSU lidar sodium density contour plot. Scale is m^{-3}. This data is for one telescope, and from the night of UT day 289 of 2009.

One obvious use of the lidar Na density is to verify the existence of the MLT sodium layer and nightglow production. The lidar-reported density is the density of ground state sodium. Both ground state sodium and O_3 are required to begin the Chapman mechanism in (R1). As mentioned in section 2.4, however, atomic oxygen is required to complete the process, and if there is no atomic oxygen, then NaO removal

into stable reservoir species will dominate. The presence of this ground state Na, as evidenced by the lidar results, demonstrates that the Na nightglow emission process is active (Plane, Personal communication 2010).

CHAPTER 7

SPECTROMETER DATA SETS AND ANALYSIS

The spectrometer had its first successful set of measurements on 26 September 2009. Since that date, several nights of nightglow data were collected at the CSU sodium lidar site near Fort Collins, CO (40.6^{o}N, 105.1^{o}W). There were two modes of operation for the instrument: stand-alone, and with concurrent Na lidar observations. This chapter will describe the data collection and data analysis building on the discussion of chapter 5, and provide the FFBS results.

7.1 Data collection specifics

While the spectrometer was designed for a vertically-pointing, 14" Celestron Schmidt-Cassegrain telescope, at times it became necessary to use the Celestron pointed 30^{o} off zenith to the north, or a vertically-pointing, 75 cm diameter Newtonian telescope pointed vertically. The signal levels for all of these modes are similar. This is true even though the fiber coupling optics of the Newtonian telescope are those optimized for the lidar, not those optimized for the spectrometer (see Acott (2009) for a description of the fiber coupling optics for the Newtonian telescope).

The moon could cause a problem for the results; therefore, the moon position is tracked using the MyStars! computer program to determine if the moon passes within or

near the spectrometer field of view. For the most part, a moon far from the field of view appears to have little to no effect over the spectrometer's ability to collect useful nightglow data. This can be seen by comparing data for nights where there was a moon for only part of the evening (such as 26 September 2009). By noting when the moon set and where it was located in the sky. No apparent dependence on the location of the moon can be seen in the data. However, measurements on other nights where the moon passes near the spectrometer field of view were made using a 4-position spectrometer mode: D1, D2, dark counts, and the 532 nm interference filters, with the Faraday filters in the beam path. The signal through the 532 nm interference filter does show some effects from the moonlight; when comparing the path of the moon across the sky to these background signal levels, similar trends are noted. Therefore, on nights of very bright moons it may be necessary to subtract off this background from the nightglow signals.

7.2 Data simulation

In order to determine the dependence of the FFBS results on knowledge of γ and the *c*-coefficients as well as other factors, sets of simulated data were created. These were created by assuming a level of nightglow of 150 Rayleigh and a telescope half field of view of 0.089° and 0.3% transmission by the telescope, fiber coupling, interference filter, passive Faraday filter losses, and PMT quantum efficiency. This transmission is much less than would be expected based on an analysis of the properties of the optics; however, 0.3% is what is required to create signal levels similar to those measured by the spectrometer. The source of this discrepancy is not clear. The equations of chapter 5 were used to calculate the four *S* values from a particular values of R_{DA}, R_{DX} and $R_{D.}$ To

determine sensitivities Tables 7.1 and 7.2 show their fractional change for perturbations of each parameter in turn. More importantly, it also details the changes in calculated ratios when different values of the nightglow spectra (leading to different *c* coefficients) are assumed, as discussed in chapter 3. For Table 7.1 the data were generated using the nightglow spectrum for a NaO bond energy of 21219 cm^{-1}, a set of experimentally measured Faraday filter functions, and a γ value of 0.63. The same values were used for table 7.2 except that the experimental Faraday filter scans were replaced with theoretical Faraday filter functions using the theoretically-optimum temperatures of 459 K and 454 K for filter A cell and tip, respectively, and 450 K and 445 K for filter B cell and tip, respectively, to generate the data.

The changes in R_D, for the most part, are relatively small, within 5% of the initial value. For many of the changes, such as assuming a mean LOS wind or different NaO + O Boltzmann distribution temperature (see section 3.6.2), there is no change at all, meaning that these factors can be ignored. However, while the measurements for R_D are quite robust, measurements of the other parameters are much more sensitive to changes in the *c* and γ parameters, meaning good measurements of *c* and γ are necessary for accurate determination of R_{D1}, R_{D2}, R_{DX}, and R_{DA}. The changes due to Faraday filter function variation are somewhat larger, meaning that it is important to have accurate measurements of the Faraday filter transmission functions at a given time, as the Faraday filters do vary over long time scales of a month or longer, as demonstrated in Figure 7.1. The change in temperature in Table 7.2 roughly approximates the changes that are seen (about 0.05 change in transmission at line center) in the actual Faraday filter functions over four months. Such changes are due to imperfections in the thermal isolation of the

unit. Thus, the Faraday filter transmission should be measured from time to time as needed.

Table 7.1: Results of the FFBS data simulation. For spectra “original” indicates the 21219 cm¬-1 NaO bond energy derived spectra, including one vibrational excitation for both A and X reaction pathways. Unless otherwise indicated, the temperature of the COM of the NaO + O system used is 200 K, and the line of sight wind is 0 m/s. “Sipler and Biondi” refers to using the upper and lower limits as found in their 1978 paper, namely a Na atom dissociation kinetic energy of 282 cm^{-1} for reaction R2, and 403 cm^{-1} for reaction R2b. The data were generated using R_D=1.7, R_{D2}=2.0, and R_{D1}=1.5; which gives a value of R_{DA}=1.89 and R_{DX}=1.42.

γ used in analysis	spectra used in analysis	NaO_2 temperature used to generate data	LOS wind used to generate data	fractional change in R_D	fractional change in R_{DA}	fractional change in R_{DX}	fractional change in R_{D2}	fractional change in R_{D1}
0.64	Original (see table heading)			-0.01	-0.26	0.85	-0.19	1.01
0.62	Original			0.01	0.46	-0.37	0.24	-0.47
0.63	No O_2 excitation			0.01	-0.61	-0.14	-1.08	-1.22
0.63	O_2: R2b=0 R2=1			0.01	-0.62	-0.26	-1.23	-1.44
0.63	O_2: R2b=0 R2=2			0.01	-0.50	0.63	-0.84	-0.47
0.63	O_2: R2b=0 R2=3			-0.01	-0.41	0.26	-0.95	-0.89
0.63	O_2: R2b=1 R2=2			0.01	0.46	1.52	-1.68	-2.18
0.63	O_2: R2b=1 R2=3			0.03	0.26	0.40	-1.22	-1.25
0.63	Sipler and Biondi			0.04	-0.12	0.17	-1.58	-1.77
0.63	Original	50 K		0.00	-0.03	0.06	-0.16	-0.08
0.63	Original		50 m/s	0.01	0.10	-0.14	0.11	-0.13
0.63	Original		-50 m/s	-0.01	-0.09	0.17	-0.09	0.16

Table 7.2: Results of the FFBS data simulation. Data generated using the theoretically-generated Faraday filter functions as described in the text and analysis Faraday filter theoretical temperatures given in the table. Other generation and analysis parameters were the same as described in the text and in the heading for Table 7.1.

	Faraday filter A Cell temperature (K)	Faraday filter A Tip temperature (K)	Faraday filter B Cell temperature (K)	Faraday filter B Tip temperature (K)	fractional change in R_D	fractional change in R_{DA}	fractional change in R_{DX}	fractional change in R_{D2}	fractional change in R_{D1}
Data generation values	459	454	450	445					
Analysis values	459	453	450	444	-0.02	0.30	24.70	-2.5	-30.4
Analysis values	459	455	450	446	0.05	1.70	0.78	-1225	-1.15

7.3 Faraday filter stability

As mentioned to at the end of section 7.2, Faraday filter stability is important for the proper operation for the FFBS. Figure 7.1 shows a plot of the four Faraday filter transmission functions measured at various times, each at least approximately one month apart.

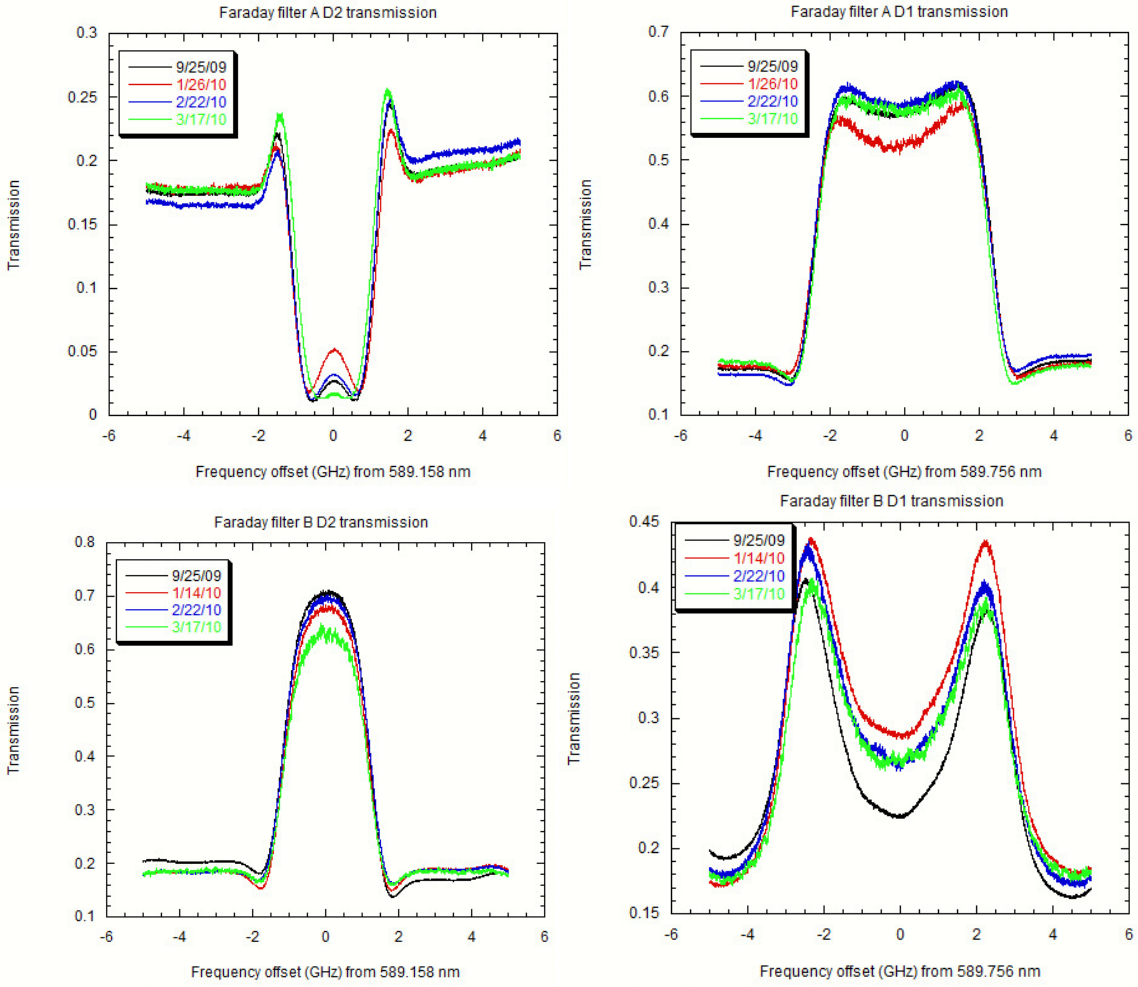

Figure 7.1: Long-term Faraday filter stability. Top Left: Faraday filter A D_2 transmission, Top Right: Faraday filter A D_1 transmission, Bottom Left: Faraday filter B D_2 transmission, Bottom Right: Faraday filter B D_1 transmission. In each plot, black is the measurement from 25 September 2009, red is from 14 January 2010, blue is from 22 February 2010, and green is from 17 March 2010. Note that Faraday filter A was dismantled before the 17 March measurement.

As the figure shows, many of the variations in Faraday filter transmission functions are quite small. Therefore, for the present construction, the Faraday filter functions only need to be updated each month to ensure proper operation and reliability of the FFBS. The difference in theoretical Faraday filter temperatures used for data generation and analysis of the data simulation in Table 7.2 are similar to the maximum difference of transmission functions seen in Figure 7.1. Therefore, even for small changes in Faraday filter transmission, the results for R_D should be quite reliable.

7.4 Data Sets

The FFBS began operation in September 2009. Since that date, five nights of moonless, cloudless nightglow data were collected before observations ceased in April 2010. Several nights have a lunar background, which has been subtracted. Two nights in April 2010 were taken with different Faraday filter temperature settings in an attempt to increase the signal to noise ratio. The following sections detail the specifics of these three data sets and particulars of the data analysis.

7.4.1 Moonless, cloudless data sets

Of the five moonless and cloudless nights, two in March 2010 had collocated lidar observations. Figures 7.2-7.6 show the results of these five nights.

Table 7.3 lists the Faraday filter c coefficients and γ values used for all of the data. Table 7.4 and 7.5 summarize these results, and contains a similar summary of ESI data from Slanger *et al*. (2005). The FFBS data compares well with the Slanger results; there are differences in the absolute values which are reasonable based on geophysical

variability, but the nocturnal variations in both sets (represented by the standard deviation in the table) are very similar. Figure 7.7 shows nightly-averaged R_D values from both the FFBS and ESI measurements. To make a more detailed comparison with the Slanger data, however, more FFBS data would be required.

Unfortunately, as the figures reveal, the error in several of the data points for R_{DX}, R_{DA}, R_{D1}, and R_{D2} are quite large, and several of the points have negative values. (The exact fraction of "good" points varies from night to night, the average fraction of good points for the five data sets and all ratios is 0.76. "Good" points are defined as those with positive mean value and error less than the mean value.) The reason for the negative values lies in the differences in the numerators of equation (5.4). The noisy nightglow signals, after correction with γ and the *c* coefficients, can cause the differences to vary from positive to negative values. Due to low and noisy signal levels, the differences appearing in the denominators of equation (5.12) and (5.13), (such as $\frac{c_{B2X}}{\gamma} S_{A2} - c_{A2X} S_{B2}$) can be very small and even negative. This causes the calculated error in $R_{D(A)}$, $R_{D(X)}$, R_{D1}, and R_{D2} to increase to unexpectedly large levels at some observation intervals. This problem is eliminated in the R_D calculation, since the differences of the *c* coefficients multiplied by the signals decreases the contribution of the noisiest signals as can be seen in (5.8) and (5.15).

Table 7.3: Faraday filter *c* coefficients and γ values for data presented in Section 7.4.1

Date range		c_{A1}	c_{A2}	c_{B1}	c_{B2}	γ
9/26/09	NaO(A)	0.5068	0.1199	0.2798	0.4145	0.731
	NaO(X)	0.5542	0.0982	0.2724	0.4777	
9/30/09	NaO(A)	0.5068	0.1199	0.2798	0.4145	0.351
	NaO(X)	0.5542	0.0982	0.2724	0.4777	
10/7/09	NaO(A)	0.5068	0.1199	0.2798	0.4145	0.239
	NaO(X)	0.5542	0.0982	0.2724	0.4777	
2/23/10	NaO(A)	0.5157	0.1205	0.3136	0.4104	0.540
	NaO(X)	0.5645	0.0991	0.3102	0.4718	
3/15/10 and 3/16/10	NaO(A)	0.4994	0.1289	0.3037	0.3775	0.290
	NaO(X)	0.5484	0.1067	0.3039	0.4313	
4/19/10	NaO(A)	0.4193	0.2068	0.2984	0.4201	0.472
	NaO(X)	0.4385	0.0230	0.2935	0.4828	

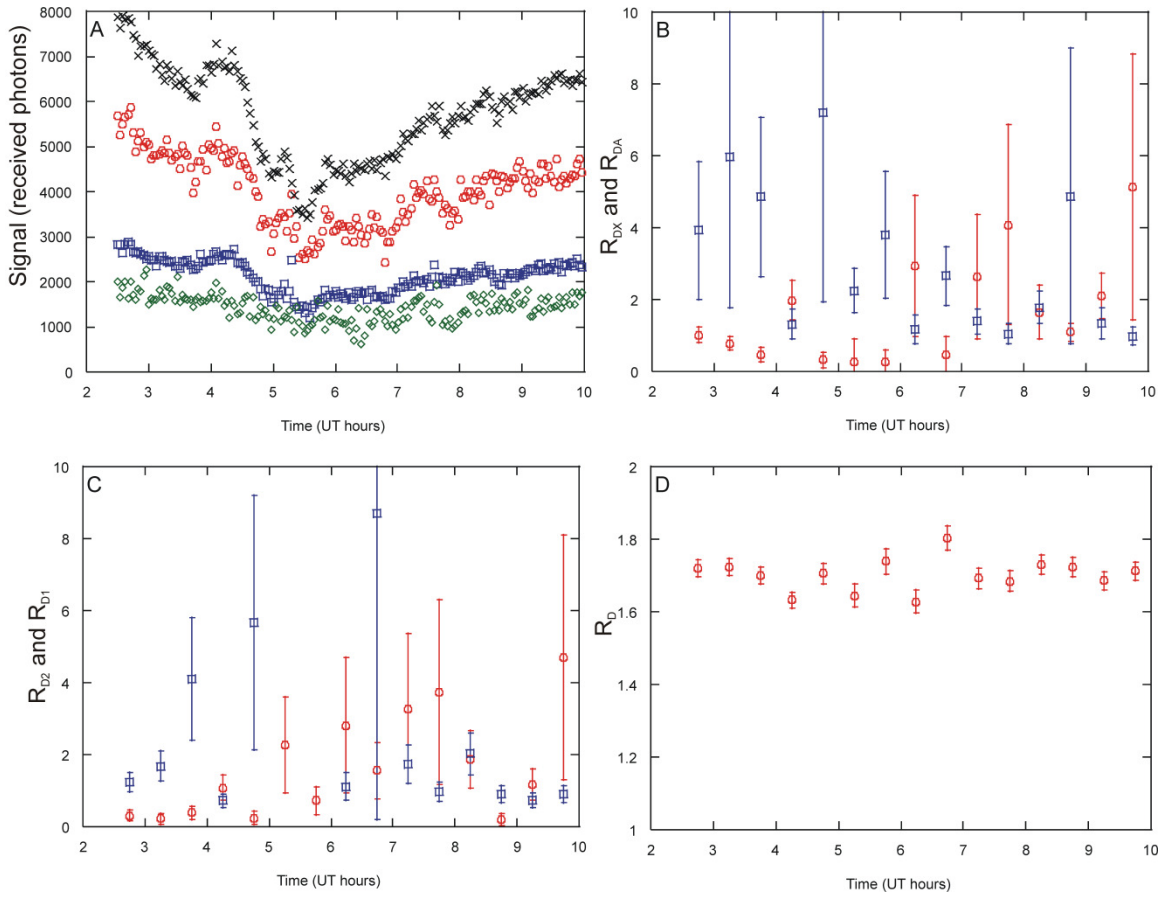

Figure 7.2: Nightglow data from 26 September 2009. (A) Raw signal with dark counts removed. Red: Faraday filter A D_1 signals (corrected by γ), blue: Faraday filter B D_1 signals, green: Faraday filter A D_2 signals (corrected by γ), and black: Faraday filter B D_2 signals. (B) R_{DX} (red) and R_{DA} (blue) measurements. (C) R_{D1} (red) and R_{D2} (blue) measurements integrated for 30 minutes. (D) R_D measurements. (B) through (D) have been integrated for 30 minutes and error bars due to photon noise.

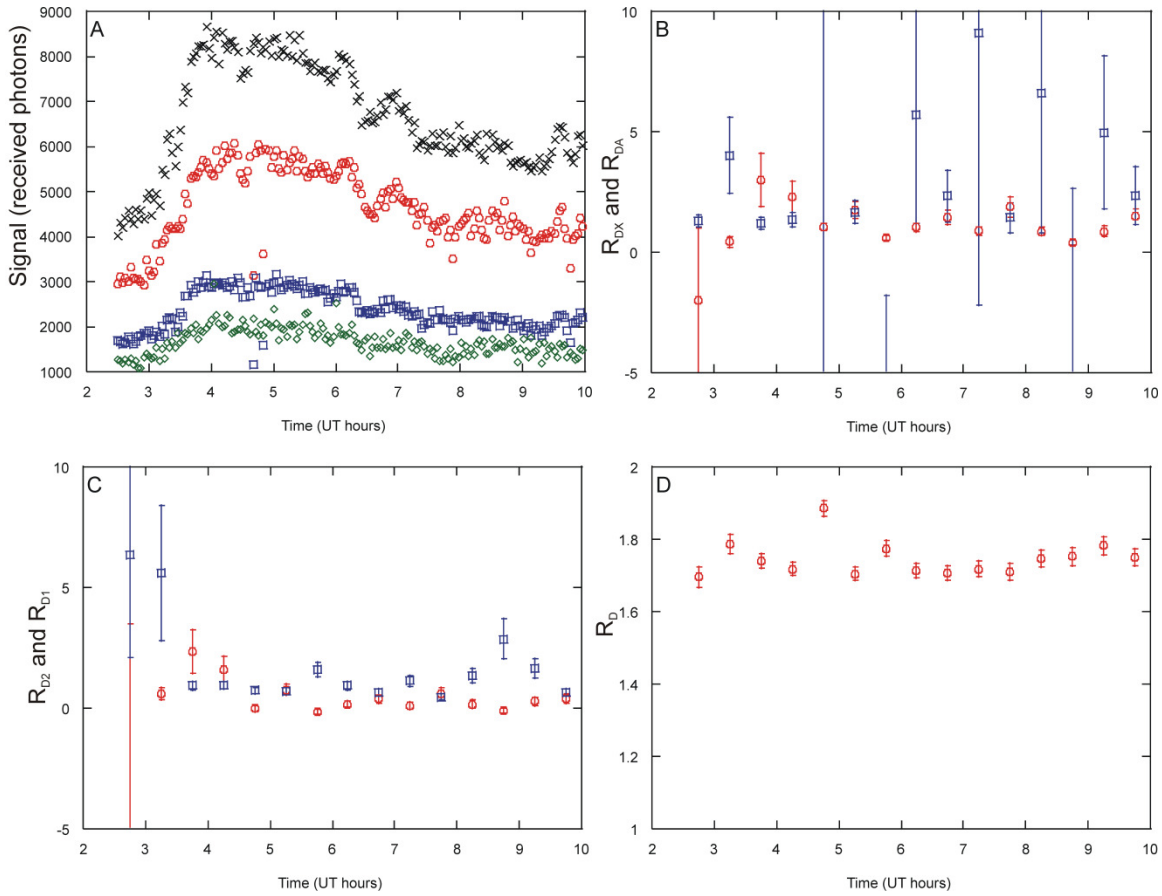

Figure 7.3: Same as figure 7.2, except for 30 September 2009.

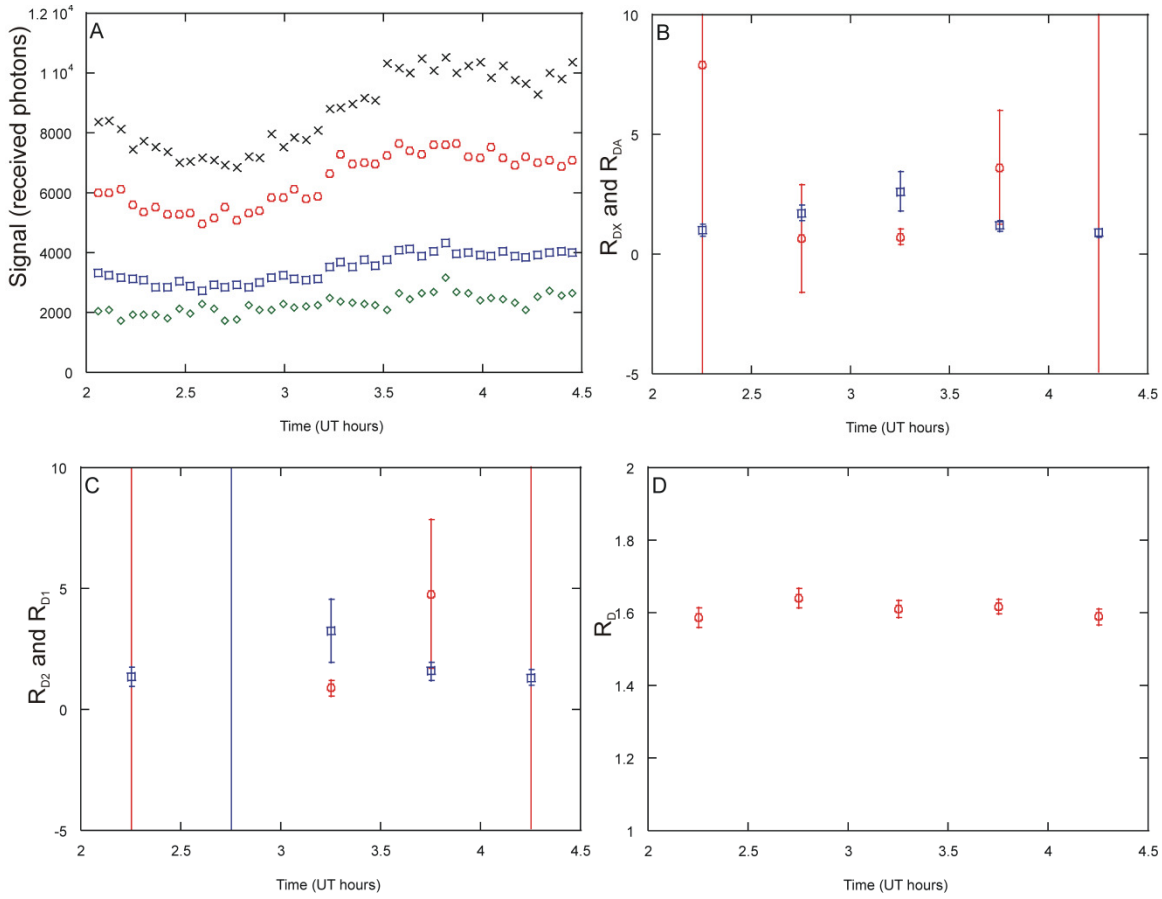

Figure 7.4: Same as figure 7.2, except for 7 October 2009.

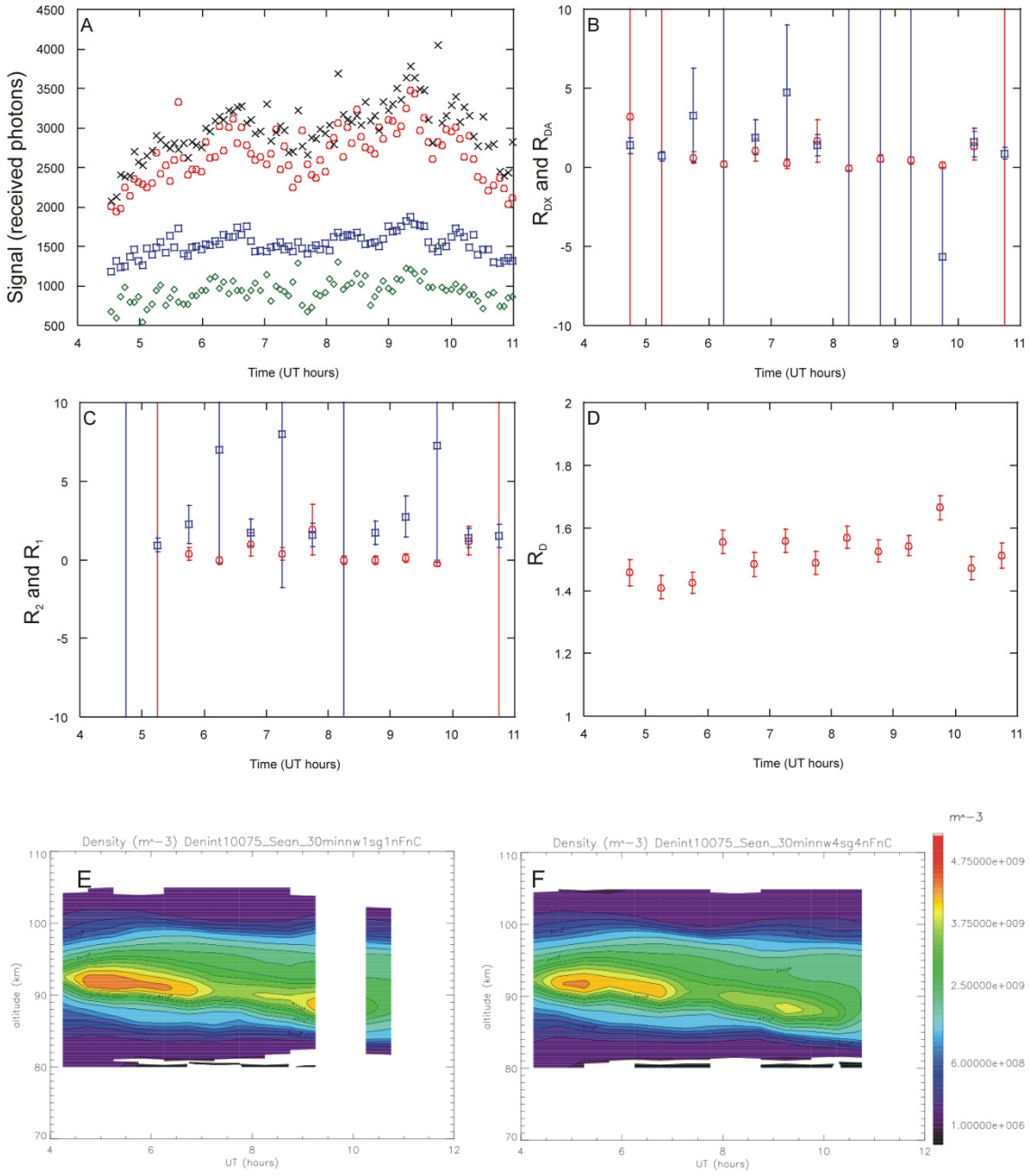

Figure 7.5: Same as figure 7.2, except for 15 March 2010. (E) and (F) are lidar measured sodium density contour plots for 30° off zenith to the north and 20° off zenith to the east, respectively.

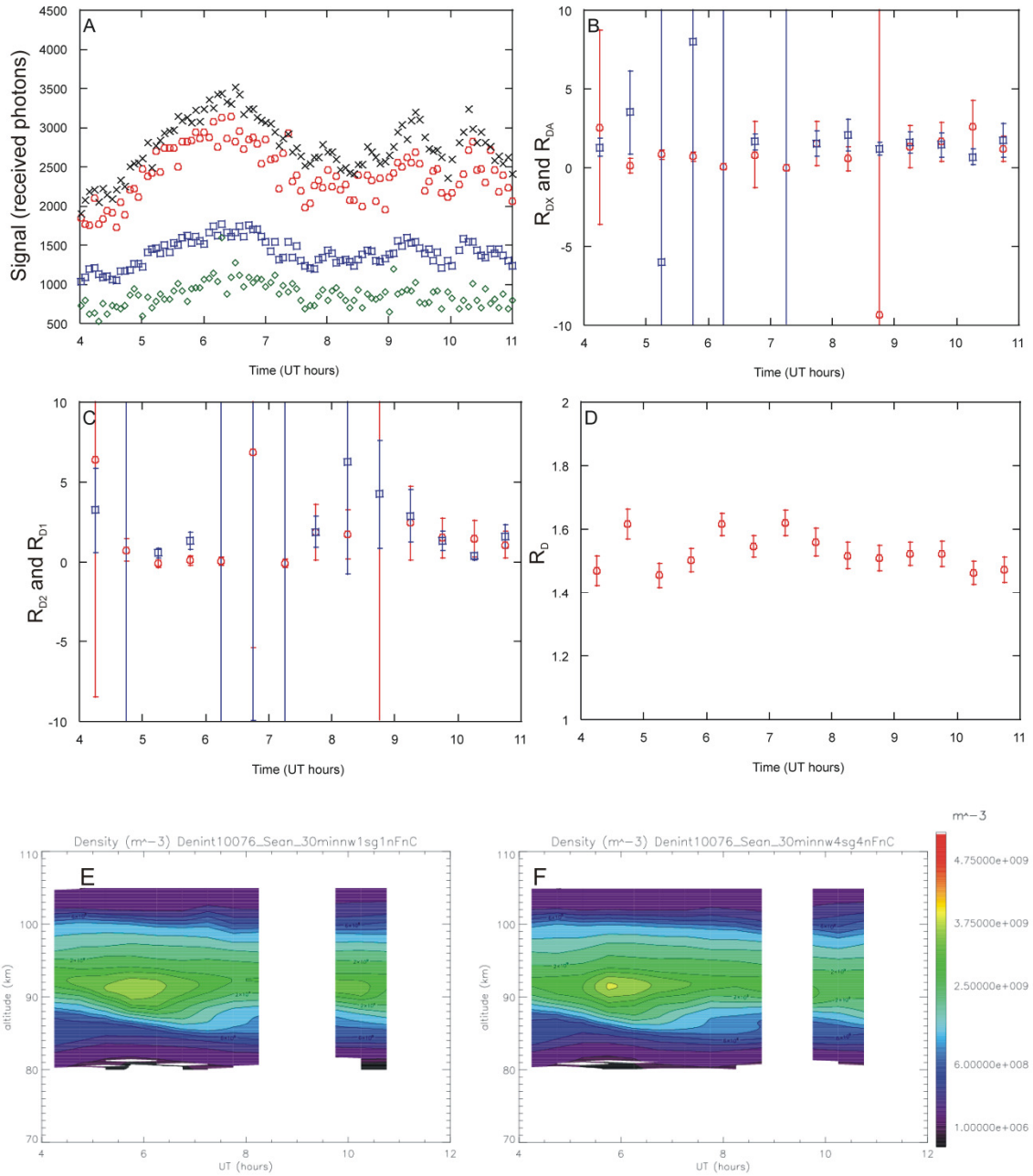

Figure 7.6: Same as figure 7.5, except for 16 March 2010.

Table 7.4: Summary of data and comparison to Slanger et al (2005).

FFBS results			Slanger et al Results		
Date	Duration (hours)	R_D and SD	Date	Duration[a]	R_D and SD
9/26/09	7.50	1.70±0.04			
9/30/09	7.50	1.74±0.05			
10/7/09	2.27	1.61±0.02			
			10/21/00	11	1.582±0.049
			10/22/00	5	1.604±0.046
			10/23/00	8	1.664±0.020
			10/24/00	9	1.613±0.034
			10/25/00	11	1.563±0.027
			3/1/00	8	1.582±0.034
			3/2/00	5	1.569±0.066
			3/3/00	9	1.568±0.059
			3/4/00	13	1.651±0.021
			3/5/00	9	1.594±0.045
3/15/10	6.50	1.51±0.07			
3/16/10	6.50	1.53±0.05			

a. The data from Slanger et al. is given in terms of number of spectra for each night; each spectrum ranges in duration from 40-70 minutes.

Table 7.5: Observed nightly mean of $R_{D(A)}$, $R_{D(X)}$, R_{D1}, and R_{D2} results The number in parentheses next to each average is the number of points used to determine the average. The number in parentheses next to the duration is the total number of data points.

Date	Duration (hours)	$R_{D(X)}$ and SD	$R_{D(A)}$ and SD	R_{D2} and SD	R_{D1} and SD
9/26/09	7.50 (15)	1.67±1.48 (15)	2.97±2.01(15)	2.35±2.41(13)	1.64±1.44 (15)
9/30/09	7.50 (15)	1.29±0.74 (14)	3.0±1.98 (11)	1.77±1.81 (15)	0.74±0.70 (10)
10/7/09	2.27 (5)	2.17±2.05 (2)	1.49±0.71 (5)	1.88±0.92 (4)	1.71±1.95 (2)
3/15/10	6.50 (13)	0.96±0.48 (6)	1.98±1.36 (8)	1.76±0.55 (8)	1.00±0.62 (5)
3/16/10	6.50 (14)	1.13±0.40 (6)	1.68±0.74 (10)	1.95±1.27 (9)	1.56±0.54 (7)

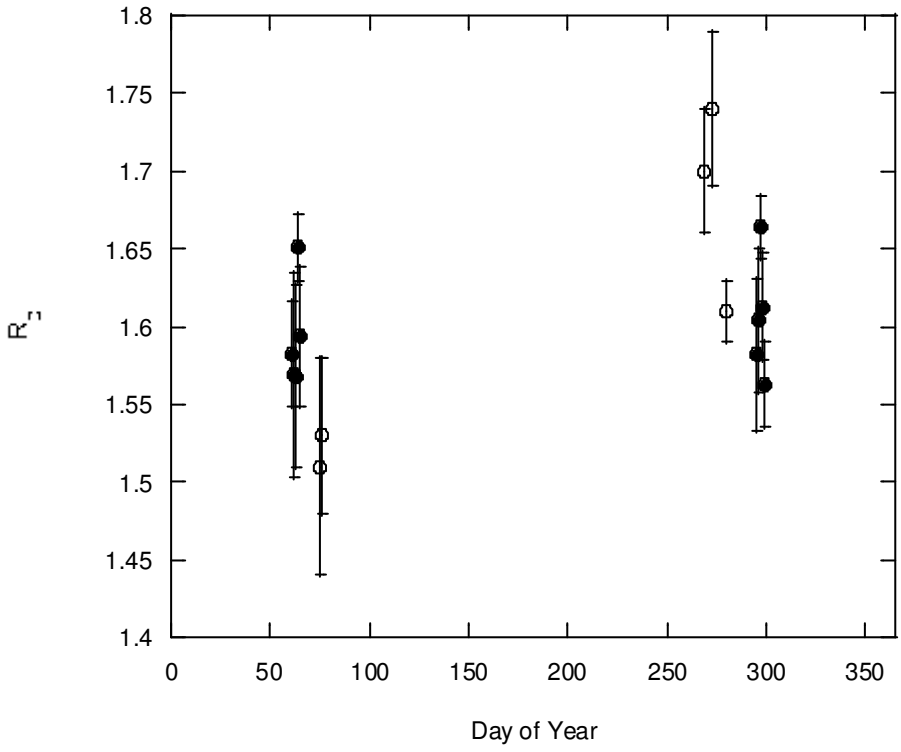

Figure 7.7: Plot of the nightly average R_D values from Table 7.4, as a function of the day of the year. Solid points are the data from Slanger et al (2005), and open points are the FFBS results. The error bars in both cases are equal to one standard deviation.

As can be seen in Figures 7.5 and 7.6, the lidar sodium density and the nightglow signal variations due appear to track each other (however there are differences in the variation of the individual signals which is expected because of the modified Chapman mechanism). Temporal variations in R_D throughout the night are a bit more difficult to compare; more data would be necessary to make an accurate comparison. However, it generally appears that as the sodium layer moves to a lower altitude, R_D increases. Measurements of $[O_3]$, [O], and $[O_2]$ (from TIMED/SABER overpass data, for example) would also help in the understanding of these variations. These quantities would help to compare the FFBS results with the expectations of the modified Chapman mechanism.

Table 7.5 represents the first measurements of these quantities based on the modified Chapman mechanism. From the laboratory experiments from Slanger et al (2005), it is expected that R_{DA} should be greater than R_{DX}, which appears to be the case for these nightly averages. The one exception is 7 October 2009, but this data collection period was extremely short. The expected values of R_{D2} and R_{D1} are not known; specific knowledge of the reaction rate constants and the fractional production of $^2P_{1/2}$ and $^2P_{3/2}$ Na states for each NaO reaction would be required in order to draw conclusions from the FFBS results.

7.4.2 Data with lunar background

Several data sets were collected on nights where the moon passed very close to the spectrometer field of view. Clearly, this could skew the FFBS results. Luckily, as can be seen in figure 7.8, the 532 nm interference filter signals show a dependence on the lunar motion as shown in the MyStars! program, but this 532 nm measurement shows

little of the variation present in the nightglow channel in the second half of the night as the moon approaches the horizon. For nights where there is no moon, the 532 nm interference filter signals are very similar to the dark count measurements. Therefore, I will assume the 532 nm interference filter signals ($s_{532A,B}$) can be used to subtract off the contaminating lunar signal in the nightglow channel.

However, there is an unknown scale factor to convert the 532 nm filter signals into the proper scale for subtraction from the nightglow measurement channels. An approach is to fit the nightglow signals (including a lunar background and dark counts) to a fifth-order polynomial plus the moon background channel:

$$s_{fit} = A + Bt + Ct^2 + Dt^3 + Et^4 + Ft^5 + Gs_{532A,B}(t) \tag{7.1}$$

where the parameters *A* through *G* are determined by the fitting procedure. The fit is achieved with the SVDFIT procedure in IDL (based on the linear least-squares fit by singular value decomposition process described in *Numerical Recipies in C* (Press, et al. 1992)). This fit is used to determine the value of *G*, which represents the strength of the lunar signal that is present in the nightglow signals. Once the value of *G* for all four nightglow signals has been determined, $Gs_{532A,B}$ can be subtracted from each actual measured signal. Then, data processing can proceed as described in chapter 5.

To ensure that the lunar variation is well represented for subtraction, I first smooth the 532 nm filter signals with a 7-point moving boxcar average. Since almost all of the nights with the moon present within the field of view have data during times when the moon is well outside of the field of view, or not present at all, then the minimum value of the smoothed data set is subtracted from the data set. Then the lunar signal is normalized to the maximum value. The result is a curve, ranging from zero at the

minimum lunar signal to one at the maximum lunar signal. The zero-point minimum ensures that little to no subtraction will occur when the lunar signal is not present in the nightglow data signals.

Figures 7.8 through 7.10 show the results of the lunar signal subtraction for a representative night (23 February 2010). Figure 7.8 shows the two lunar signals as measured with the 532 nm interference filter in the FFBS and the zero-point subtracted and normalized curve used for the fitting procedure. Figure 7.9 compares the nightglow signals before and after the lunar signal subtraction. Figure 7.10 shows the analysis products, with a comparison of the R_D results with and without the lunar signal subtraction. Notice in Figures 7.9 and 7.10 that there is little change in the values during the second half of the observation period; this is expected since the moon is below the horizon by UT hour 10. In figure 7.10, the error bars have been extended by adding the square roots of the nightglow signal, the dark count, and the original 532 nm interference filter signal in quadrature for each time interval.

The results for R_D for lunar subtraction show a decrease in the value before 7 UT, when compared with the data with no lunar subtraction. The data sets begin to converge after 7 UT, becoming almost identical at 10 UT. This is to be expected, as the moon is closest to zenith at 3 UT, as can be seen in Figure 7.9. By 10 UT, the moon is below the horizon, and so there should be little subtraction at this time.

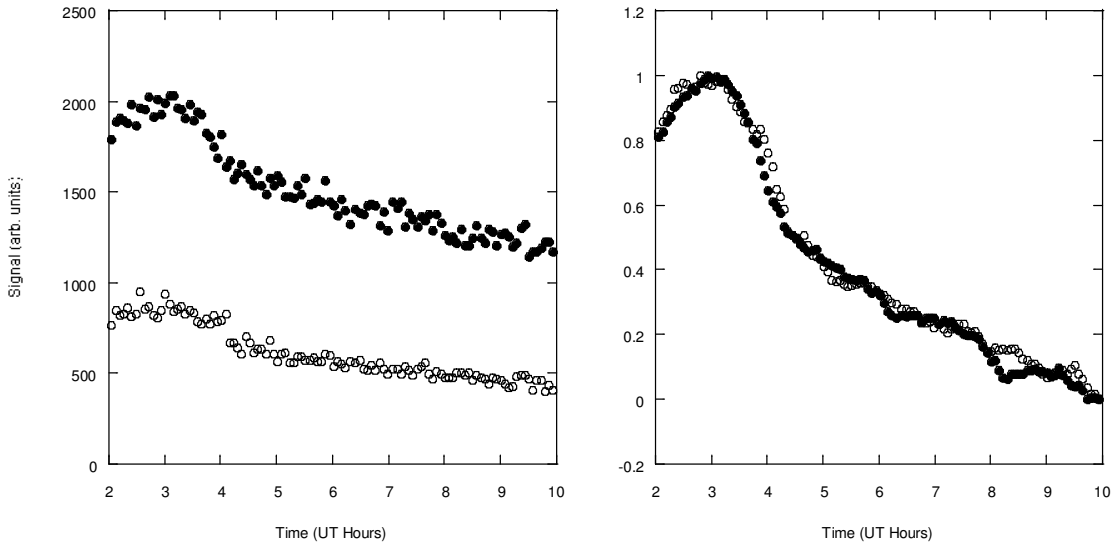

Figure 7.8: Lunar signal for 23 February 2010. Left panel: raw signal from the 532 nm interference filter for Faraday filter A (open points) and Faraday filter B (solid points). Right panel: Smoothed, zero-point subtracted, and normalized signals, points are the same as the left panel.

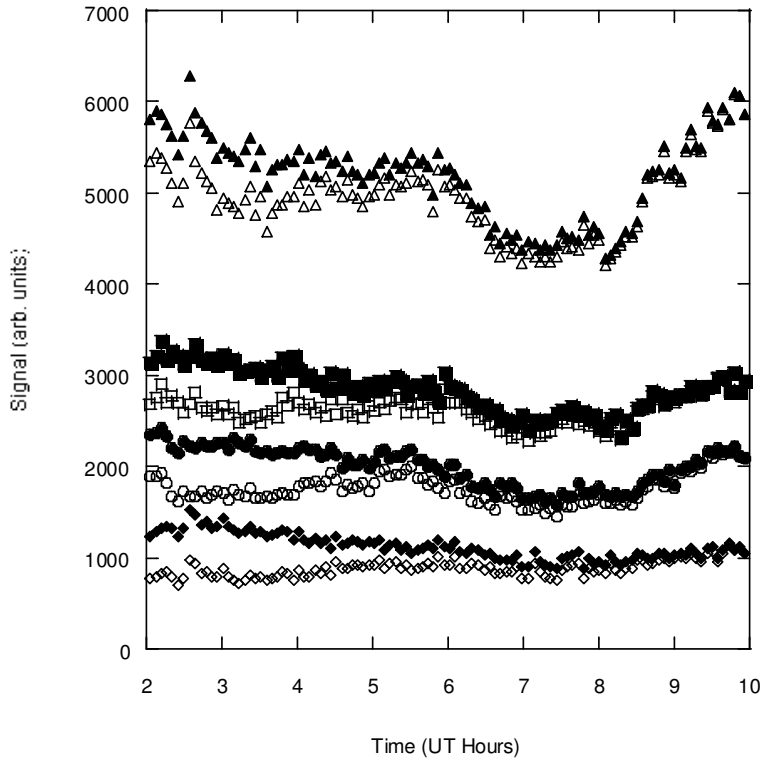

Figure 7.9: Comparison of nightglow signals with and without lunar signal subtraction for 23 February 2010. Note that there is no dark count subtraction. Solid points are with no lunar signal subtraction, open points are after lunar subtraction. The sets of curves are as follows from top to bottom: D_2 signal for Faraday filter B, D_1 signals for Faraday filter B, D_1 signals for Faraday filter A, and D_2 signals for Faraday filter A.

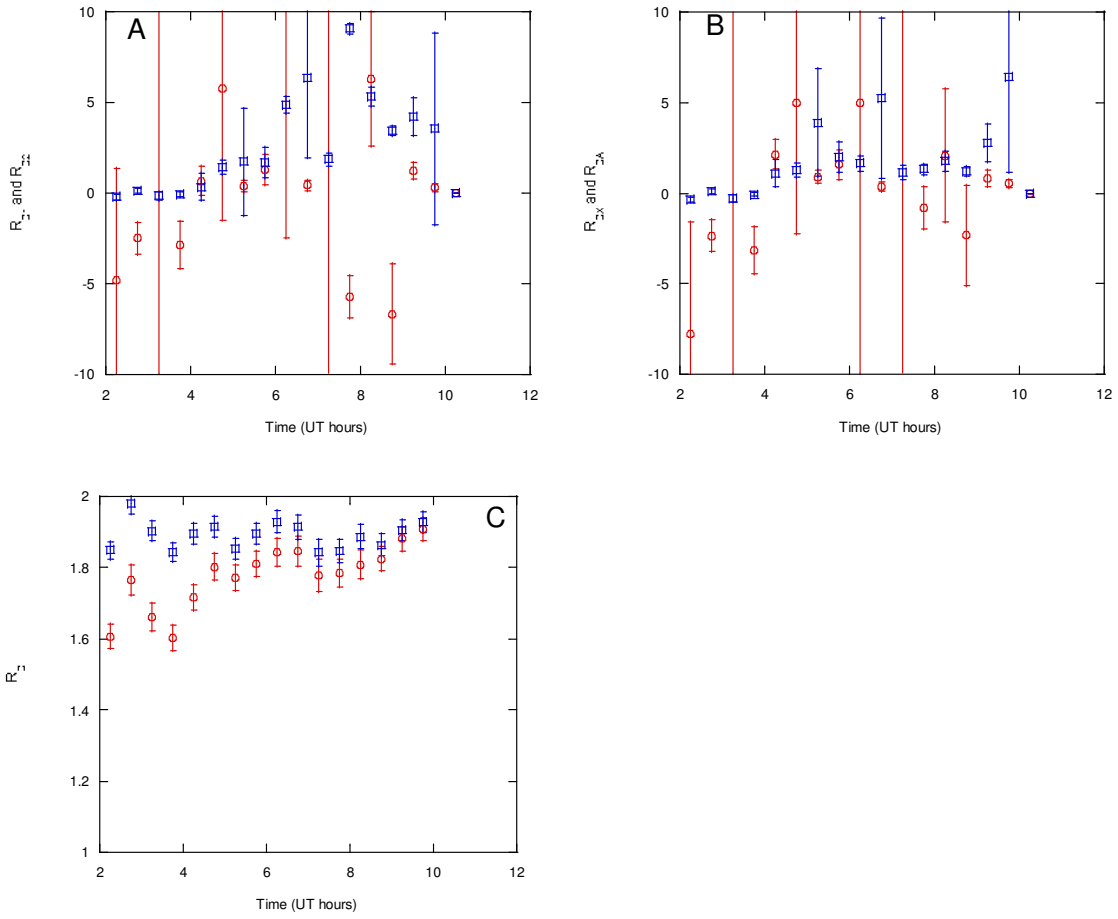

Figure 7.10: Analyzed results from 23 February 2010. (A) R_{D1} (red) and R_{D2} (blue). (B) R_{DX} (red) and R_{DA} (blue). (C) R_D for lunar signal subtracted data (red) and for no subtraction (blue).

It is important to note that the ability of the FFBS to collect data in the presence of a lunar background makes it unique among passive nightglow instruments. This ability is probably due to both the narrow field of view and the ultranarrow bandwidth of the FFBS. Other nightglow instruments, especially all-sky imagers, are unable to make measurements while the moon is high in the field of view. Others, such as the Yucca Ridge all-sky OH imager located near Fort Collins, is unable to record data at all while the moon is above the horizon (Yue, Personal Communication 2010).

7.4.3 April 2010 data with different Faraday filter temperature settings

Due to the Faraday filter temperature settings, S_{A2} has an almost zero transmission near line center (since it is maximized for transmission at the D_1 transition), leading to very low signal levels for this measurement. It was believed that by changing the temperature setting, this transmission could be increased and thus would increase the measurement precision of the derived FFBS quantities. Figure 7.11 shows the Faraday filter A functions used for April 2010 measurements (Faraday Filter B remained unchanged from the previous settings). The Faraday filter c coefficients are presented in Table 7.3. Notice that c_{A2X} and c_{A2A} are larger than the other values presented in Table 7.3.

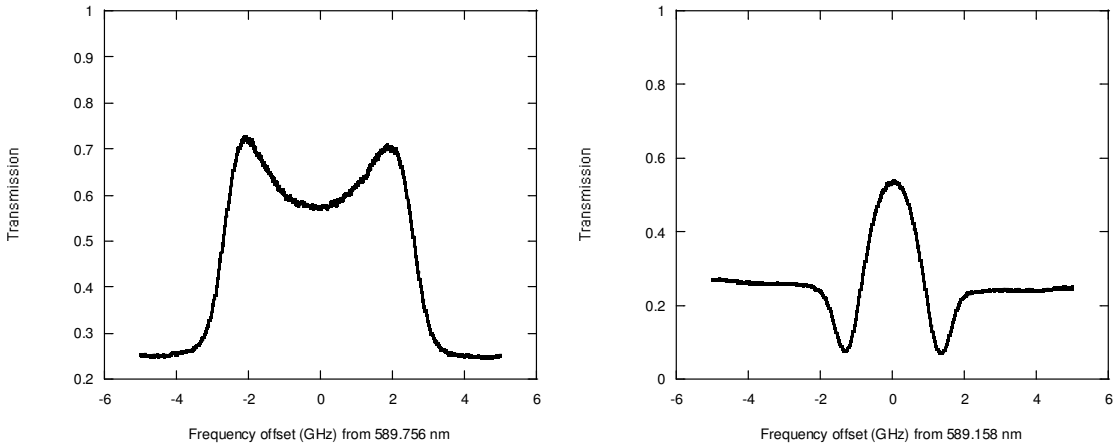

Figure 7.11: Faraday filter A functions from April 2010. Left panel: D_1 transmission, Right panel: D_2 transmission.

Two nights of data were collected under moonless and cloudless conditions. The results of one representative night are presented in Figure 7.12. The other night has similar results. Notice that the while the error bars in R_{D1} and R_{D2} are generally smaller, the error bars for many of the quantities, including R_D, are larger than the data presented in Figures 7.2-7.6.

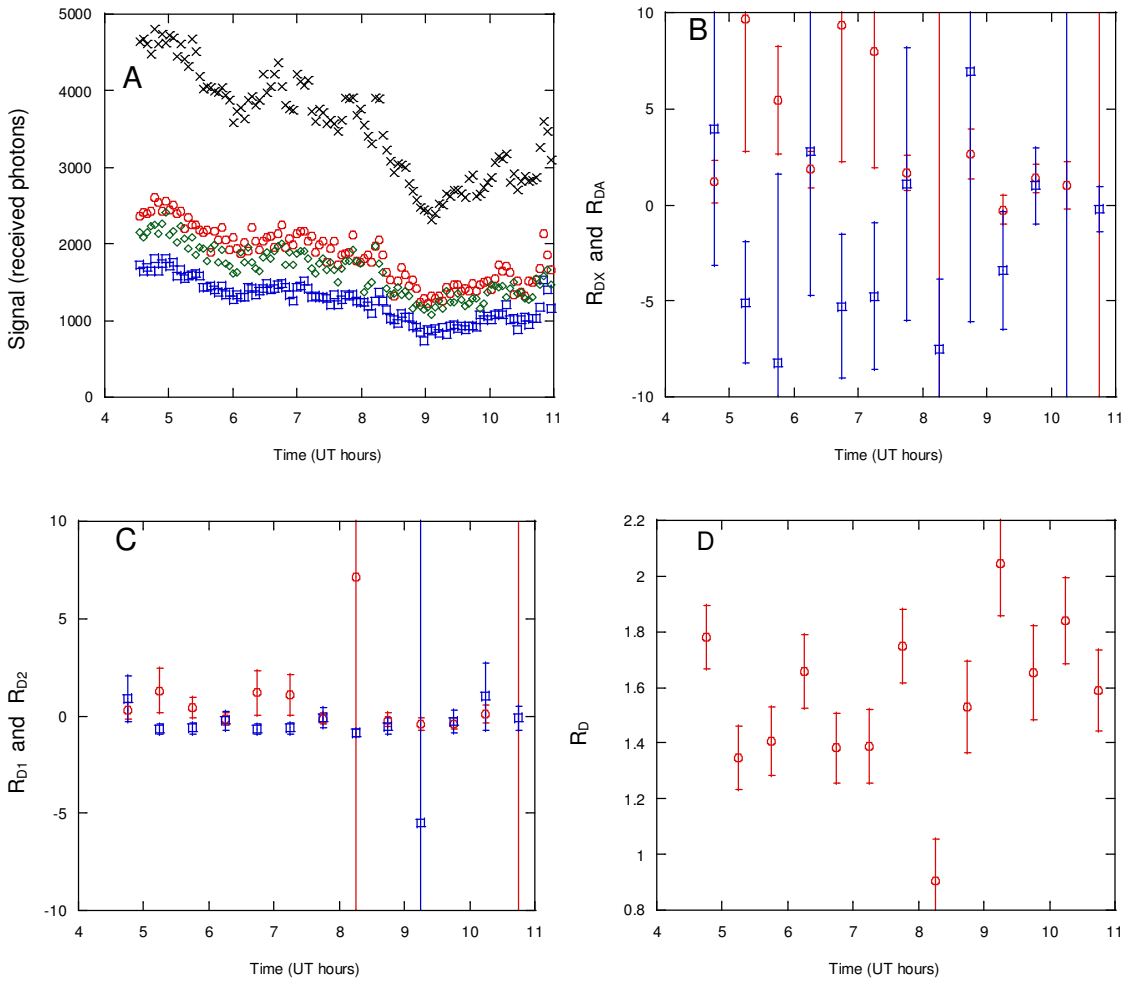

Figure 7.12: Data from 19 April 2010. (A) Raw signal with dark counts removed. Red: Faraday filter A D_1 signals (corrected by γ), blue: Faraday filter B D_1 signals, green: Faraday filter A D_2 signals (corrected by γ), and black: Faraday filter B D_2 signals. (B) R_{DX} (red) and R_{DA} (blue) measurements. (C) R_{D1} (red) and R_{D2} (blue) measurements integrated for 30 minutes. (D) R_D measurements. (B) through (D) have are integrated for 30 minutes and error bars due to photon noise.

The larger variations in values and larger error bars can be explained by inserting the coefficient values in equation (5.7), which is repeated here:

$$R_D = \frac{I_2}{I_1} = \frac{(c_{A2A} - c_{A2X})S_{B2} + \frac{(c_{B2X} - c_{B2A})}{\gamma} S_{A2}}{(c_{A1X} - c_{A1A})S_{B1} + \frac{(c_{B1A} - c_{B1X})}{\gamma} S_{A1}} \frac{c_{A1X}c_{B1A} - c_{B1X}c_{A1A}}{c_{A2A}c_{B2X} - c_{B2A}c_{A2X}} \tag{7.2}$$

By rewriting this equation with specific *c* values from Table 7.3, the increase in error shown in figure 7.8 can be explained. For reference, the average dark count subtracted signal levels for 16 March 2010 are S_{A1}=670, S_{B1}=1332, S_{A2}=238, and S_{B2}=2642. Using the *c* coefficients for March 2010 (Table 7.3), equation (7.2) becomes:

$$R_D = \frac{I_2}{I_1} = \frac{(0.0222)S_{B2} + \frac{(0.0538)}{\gamma} S_{A2}}{(0.0490)S_{B1} + \frac{(0.0008)}{\gamma} S_{A1}} 1.0009 \quad . \tag{7.3}$$

For April 2010, the coefficients from Table 7.3 yields:

$$R_D = \frac{I_2}{I_1} = \frac{(-0.0229)S_{B2} + \frac{(0.0627)}{\gamma} S_{A2}}{(0.0192)S_{B1} + \frac{(0.0049)}{\gamma} S_{A1}} 2.3263 \quad . \tag{7.4}$$

While the values are quite similar, there is one important difference. The negative coefficient difference in equation (7.4) means that the numerator can oscillate greatly, while this will not happen in equation (7.3). For reference, the signal levels on 19 April 2010 are =873, S_{B1}=1255, S_{A2}=775, and S_{B2}=3543. Thus, it is quite important to judiciously choose Faraday filter temperature settings to avoid this problem. An arbitrary change in Faraday transmission function to increase the signal level of one channel apparently does not work, and an optimization procedure to account for the signal levels is necessary for FFBS design.

7.4.4 Variation of γ and technical issues

One other issue with data collection was a film that developed on the surface of Faraday filter A's linear polarizer. This caused γ to decrease greatly, meaning that a measurement of γ was generally only valid for 48 hours. Thus, the data sets presented in

detail contain only data where the γ value was measured with confidence within 48 hours either before or after the nightglow measurement. This problem was eliminated after a thorough cleaning of the inside of the Faraday filter, and the data from after 17 March 2010 does not suffer from this problem.

For nights without valid γ measurements, it was often possible to perform an interpolation of the proper value of γ for that night by interpolating from measured γ before and after the night of nightglow measurements. The apparent appropriate curve fit is a power curve of the form $\gamma = Ax^B$, where A and B are the parameters of the fit.

7.5 Conclusions

Figure 7.13 shows all of the nightly-averaged R_D results from September 2009 through April 2010. This plot includes the moonless and cloudless data with measured γ, moonless and cloudless measurements with interpolated γ, data with the lunar subtraction, and results from April 2010 with increased error. Unfortunately, technical difficulties with the FFBS instrument prevented valid data collection during the winter months. Initial signal levels recorded suggest that the nightglow intensity in the winter should still be high enough to permit measurements with the FFBS.

A comment can be made about the ratios of R_{D2}/R_{D1} and R_{DA}/R_{DX}. These ratios are mathematically identical in terms of the theoretical terms presented in Chapter 3. If these ratios of ratios were calculated from a data set with no fluctuations, then these values would indeed be identical; however variations may cause the mean values to vary within the error bars. An example from the nightly average values from 26 September 2009 shows this is the case: R_{D2}/R_{D1}=1.43±1.35, while R_{DA}/R_{DX}=1.78±1.11. When the

error bars of the measurement are improved, the ratios R_{DA} and R_{DX} will provide information on p_{2A}/p_{1A} and p_{2X}/p_{1X}, respectively.

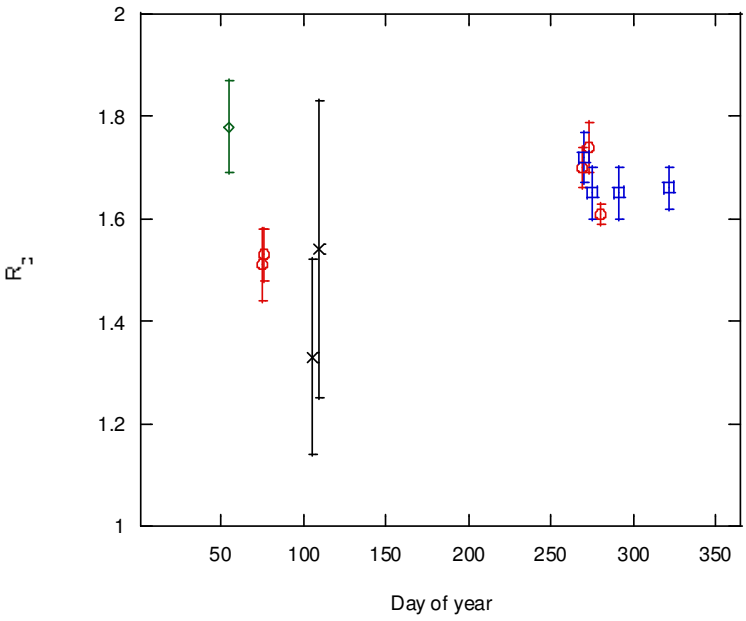

Figure 7.13: Plot of all FFBS results for R_D. Red points are moonless and cloudless data with measured γ; Blue points are moonless and cloudless data with interpolated γ; the Green point is moonlight-subtracted results; Black points are April 2010 data with increased error due to the modified *c* coefficients. The error bars are the standard deviation of the data sets.

Appendix C contains details about a different approach to data analysis utilizing curve fitting. This method has the advantage of smoothing over the short-term (less than 0.5 hour) geophysical variability and the photon noise variability. Appendix C details the initial results of performing a linear fit to the data from 26 September 2009 and suggestions for future work for utilizing this method of analyzing the data.

CHAPTER 8

CONCLUSIONS AND FUTURE WORK

This dissertation has dealt with the design and implementation of a new instrument: the Faraday filter-based spectrometer. This instrument has provided new measurements of the sodium nightglow emission in the MLT region of the atmosphere. Specifically, this instrument has provided (to my knowledge), the first measurements of R_{DX}, R_{DA}, R_{D2}, and R_{D1}. This is in addition to R_D.

The vernal and autumnal equinox FFBS measurements of R_D are similar to those of Slanger et al (2005), with values between 1.51 and 1.74. Nightly variations of R_D are also similar to those reported by Slanger et al. These first measurements of R_{DA} and R_{DX} support the assertion (based on laboratory work) that $R_{DA}>R_{DX}$. More work (both observational and laboratory) will be required to draw conclusions about the measurements of R_{D1} and R_{D2}.

8.1 Future Work

There are numerous possibilities for future work, both for continued operation of the spectrometer and improvement of the instrument.

8.1.1 Instrument Improvement

There are several ways that the instrument could be improved. Most of these are to increase the photon counts to increase the signal-to-noise ratio and decrease the error in the calculated quantities. As mentioned in chapter 5, the dominant source of error is photon noise. Other sources of error are the *c* coefficients and γ values; however these measured errors are quite small and contribute little to the overall error. Thus, increasing the signal-to-noise ratio should decrease the overall error significantly. One option is to replace the 40% quantum efficiency PMTs with higher efficiency CCD's. There are CCD devices with quantum efficiency of up to 80%, doubling the signal levels. However, a disadvantage is that the high efficiency CCDs are significantly more expensive than the 40% quantum efficiency PMTs.

Currently, the system is limited to a very small field of view because of the use of the 0.37 NA, 1.5 mm core diameter fiber. This limits the étendue of the system. Use of fiber coupling between the FFBS and the Celestron 14 allowed locating the instrument at a remote location from the telescope and, more importantly, allowed light to be collimated over 0.5 m with a spot diameter of 10 mm as required by the FFBS. It might be possible, with a carefully designed (and probably expensive) lens system to increase the field of view (and thus étendue) without using a fiber while still collimating the light with the 10 mm spot size needed to pass through the Faraday filters. Larger diameter or numerical aperture fibers would also increase the field of view, but might make the collimation through the FFBS optics more difficult.

Another option is to increase the diameter of the Faraday filters. Currently the aperture is limited by the Glan-Thompson polarizers to less than 15 mm. By increasing

the diameter, a larger beam would be able to pass through the system and might make easier the use of conventional optics to couple the FFBS and telescope and increase the field of view. There are drawbacks, however. The current magnets are not capable of producing a uniform field over a larger diameter Na vapor cell, and larger diameter polarizers would be more expensive. Temperature stability and uniformity would also be more difficult with a larger diameter Na cell.

8.1.2 Science Studies

There are several studies that could be performed with the existing data set. In addition, there is the capability of performing future collaborative studies with a variety of instruments.

While the lidar-FFBS system working in tandem can provide information on the altitude distribution of R_D, more work will need to be done to translate this data into [O]/[O_2] and investigate the variation of atomic oxygen in the MLT. Laboratory studies are ongoing (Slanger, personal communication) which will hopefully allow for the conversion of R_{D2}, R_{D1}, or R_D into [O]/[O_2]. Therefore, this dissertation is by design somewhat incomplete: the stated goal of a ground-based instrument to study the atomic oxygen variation awaits these measurements. However, the data has yielded new information about the variation of R_D and the first (to my knowledge) measurements of the individual modified Chapman mechanism pathways, the competing reactions involving NaO(A) or NaO(X).

There are at least three studies that can be performed with the current data set. One study would utilize atomic and molecular oxygen data from the SABER instrument

on the TIMED satellite as a comparison from the data from the FFBS. The SABER instrument can determine mixing ratios of O_3, O_2, and O, the three parameters needed, in addition to the Na mixing ratio, to perform a study of the modified Chapman mechanism. The SABER data could also provide some initial ability to convert our data to $[O]/[O_2]$: with luck, it will give more information on the correct "look up table" to use. SABER data at one location is limited; measurements at the same location and local time are separated by 60 days, meaning data under identical conditions as the FFBS is quite limited. SABER data profiles (limiting the profiles to between 80 and 100 km) are collected in under one minute, have a vertical field of view of 2 km, and the profiles range over 0.387° in latitude (approximately 44 km horizontal distance) and 0.284° in longitude (approximately 23 km horizontal distance). By comparison, the FFBS data as presented has been averaged temporally in half hour periods and is for one location, with a horizontal field of view at 90 km of approximately 30 m.

Another study could determine the tidal effects on the nightglow data. The lidar allows for the determination of both diurnal and semidiurnal tidal perturbations of temperature and zonal and meridional wind (Yuan, et al. 2006; Yuan, et al. 2008). The lidar can also determine the tidal perturbations of Na density. Comparing nightglow data with the lidar-determined tides will give a sense of the tidal-dependence of the Na nightglow, and hence the tidal variation of atomic and molecular oxygen.

Further collaboration could be achieved by moving the FFBS to other observing sites. For example, the Compact Echelle Spectrograph for Aeronomical Research (CESAR) (Grill, et al. 2009) is designed to measure nightglow emissions from across the entire spectrum from 300 to 1000 nm. CESAR will be located in Poker Flat, Alaska.

CESAR will provide information about the entire nightglow spectrum, including Na, OI, O_2, OH, and others. However, the resolution of this instrument is not high enough to make detailed measurements of the two nightglow production pathways. Therefore, a co-location of the FFBS would provide the ability to study in detail the Na nightglow spectrum in an environment where there is also information about other oxygen-related airglows, providing the ability to study more deeply the MLT chemistry than with either of these instruments alone.

It is important to point out how the FFBS fits into a larger picture of the work that has been ongoing for the past twenty years by the Na lidar group at CSU. The lidar system has been upgraded from a two-frequency instrument capable of determining nocturnal temperatures and Na densities, to a three-frequency system capable of determining full diurnal cycle temperatures, horizontal winds. From these measurements, diurnal and semidiurnal tides in temperature and zonal and meridional winds have been determined and compared to model climatologies; momentum flux has been calculated and used to study gravity wave breaking.

More recently, the studies have taken a slightly different direction: collaborative science with collocated instruments. This is evidenced by the FFBS as well as studies of gravity waves using the Kyoto University imager located near Fort Collins (Yue 2009). The lidar system was transferred to Utah State University (USU, approximately 42°N; similar to CSU at 41°N) on 1 April 2010. As the CSU lidar system becomes the USU lidar system, I believe that the most logical next step for science studies would be to continue these types of collocated instrument studies. USU will provide for enhanced opportunities for these types of collaborations due to the varied expertise of their faculty.

While it is sad to see the lidar system, and the FFBS, move from Fort Collins, Colorado, they should have a bright and productive future in Utah.

8.2 Final Conclusions

This dissertation has given many details about the design, theory of operation, and results from the Faraday-filter based spectrometer. Current understanding of nightglow, specifically the Na nightglow production process, and MLT region chemistry has been reviewed in order to put this instrument into context. Ancillary instruments, such as the CSU Na wind and temperature Doppler lidar, have been reviewed as well. The results of R_{DX}, R_{DA}, $R_{D1,}$ and R_{D2} are the first reported measurements of the two pathways of the modified Chapman mechanism, which previous measurements were not capable of observing; measurements of R_D verify what has been measured in the past. Future work for both improving the instrument and science studies has been suggested.

Bibliography

Acott, P. E. "Mesospheric Momentum Flux Studies over Fort Collins CO (41N, 105W)." *PhD Dissertation.* Colorado State University, 2009.

Agnelli, G., A. Cacciani, and M. Fofi. "The magneto-optical filter I." *Solar Physics* 44 (1975): 509-518.

Arimondo, E., M. Inguscio, and P. Violino, 1977. "Experimental determinations of the hyperfine structure in the alkali atoms." *Review of Modern Physics* 49, 31-75.

Arnold, Kam S., and C.-Y. She, 2003. "Metal fluorescence lidar (light detection and ranging) and the middle atmosphere." *Contemporary Physics* 44, 35-49.

Baker, D. J., and G. J. Romick, 1976. "The rayleigh: interpretation of the unit in terms of column emission rate or apparent radiance expressed in SI units." *Applied Optics* 15, 1966-1968.

Bromberg, J. Philip. *Physical Chemistry.* 2nd. Boston: Allyn and Bacon, 1984.
Chamberlain, Joseph W. *Physics of the Aurora and Airglow.* New York: Academic Press, 1961.

Chan, Y. C., and J. A. Gelbwachs,1993. "A Fraunhofer-wavelength magnetooptic atomic filter at 422.7 nm." *IEEE J. Quantum Electron.* 29, 2379-2384.

Chapman, S, 1939. "Notes on Atmospheric Sodium." *Astrophys. J.* 90, 309.

Chen, H., C.-Y. She, Paul Searcy, and Eric Korevaar, 1993. "Sodium-vapor dispersive Faraday filter." *Optics Letters* 18, 1019-1021.

Chen, H., M. A. White, David A. Krueger, and C.-Y. She, 1996. "Daytime mesopause temperature measurements with a sodium-vapor dispersive Faraday filder in a lidar receiver." *Optics Letters* 21, 1093-1095.

Clemesha, B. R., D.M. Simonich, H. Takahshi, and S. M. L. Melo, 1993. "A simultaneous measurement of the vertical profiles of sodium nightglow and atomic sodium density in the upper atmosphere." *Geophysical Research Letters* 20, 1347-1350.

Corney, A. *Atomic and Laser Spectroscopy.* Clarendon Press, 1977.

Dick, D.J., and T. M. Shay, 1991. "Ultrahigh-noise rejection." *Optics Letters* 16, 867-869.

Dressler, Edward T., Alan E. Laux, and Richard I. Billmers, 1996. "Theory and experiment for the anomalous Faraday effect in potassium." *Journal of the Optical Society of America B* 13, 1849-1858.

Edmonds, A. R. In *Angular Momentum in Quantum Mechanics*, 75-76. Princeton: Princeton University Press, 1957.

Edmonds, A. R. In *Angular Momentum in Quantum Mechanics*, 111. Princeton: Princeton University Press, 1957.

Edmonds, A. R. In *Angular Momentum in Quantum Mechanics*, 125-127, 130-132. Princeton: Princeton University Press, 1957.

Falke, S., E. Tiemann, and C. Lisdat, 2006. "Transition frequencies of the D lines of 39K, 40K, and 41K measured with a femtosecond laser frequency comb." *Physical Review A* 74, 032503.

Fricke-Begemann, Cord. *Lidar Investigations of the Mesopause Region: Temperature Structure and Variability.* Dissertation, Kuhlungsborn, Germany: Leibniz-Institut fur Atmpospharenphysik in Kuhlungsborn, 2004.

Friedman, J. S., C. A. Tepley, S. Raizada, Q. H. Zhou, J. Hedin, and R. Delgado, 2003. "Potassium Doppler-resonance lidar for the study of the mesosphere and lower thermosphere at the Arecibo Observatory." *Journal of Atmospheric and Solar Terrestrial Physics* 65, 1411-1424.

Gangrsky, Y. P., et al, 1998 "Hyperfine splitting and isotope shift in the atomic D2 line of 22,23Na and the quadrupole moment of 22Na." *The European Physical Journal, A* 3, 313-318.

Gerding, M., M. Aplers, U. von Zahn, R. J. Rollason, and J. M. C. Plane, 2000. "Atmospheric Ca and Ca+ layers: Midladitude observations and modeling." *Journal of Geophysical Research* 105, 27131-27146.

Grill, M., M. Radovan, R. Melchiorri, and T. G. Slanger. "CESAR Opto-Mechanical Design." *Eos Trans. AGU, Fall Meet. Suppl.* 2009.

Harrell, S. D., et al, 2009. "Sodium and potassium vapor Faraday filters revisited: theory and applications." *Journal of the Optical Society of America B* 26, 659-670.

Hecht, J. H., et al, 1986. "Auroral and airglow Fabry-Perot spectrometer." *Rev. Sci. Instr.* 57, 240-247.

Hernandez, G, 1975. "Reaction Broadening of the Line Profiles of Atomic Sodium in the Night Airglow." *Geophys. Res. Lett.* 2, 103-105.

Hernandez, G, 1971. "The signature profiles of O (1S) in the airglow." *Planet.Space Sci.* 19, 467-676.

Hines, C. O, 1960. "Internal Atmospheric Gravity Waves at Ionospheric Heights." *Can. J. Phys.* 38, 1441-1481.

Hoffner, J., and J. Lautenbach, 2009. "Daylight measurements of mesopause temperature and vertical wind with the mobile scanning iron lidar." *Opt. Lett.* 34, 1351-1353.

Hoffner, Josef, and Jens Lautenbach, 2009. "Daylight measurements of mesopause temperature and vertical wind with the mobile scanning iron lidar." *Optics Letters* 34, 1351-1353.

Holton, J. R, 1983. "The Influence of Gravity Wave Breaking on the General Circulation of the Middle Atmosphere." *J. Atmos. Sci.* 40, 2497-2507.

Holton, J. R., and J. Alexander, 2000. "TThe role of waves in the transport circulation of the middle atmosphere." *Geophys. Monograph* 123, 21-35.

Honig, R. E., and D. A. Kramer, 1969. "Vapor pressure data for the solid and liquid elements." *RCA Review* 30, 285-305.

Hu, Zhilin, Xianping Sun, Yiping Liu, Liping Fu, and Xizhi Zeng, 1998. "Temperature properties of Na dispersive Faraday optical filter at D1 and D2 line." *Optics Communications* 156, 289-293.

Huang, Wentao, Xinzhao Chu, B. P. Williams, S. D. Harrell, Johannes Wiig, and C.-Y. She, 2009 "Na--Edge Magneto-Optic Filter for Na Lidar Profiling of Wind and Temperature in the Lower Atmosphere." *Optics Letters* 34, 199-201.

ITT Corporation. "Interactive Data Language." Boulder, Colorado.

Jackson, John David. *Classical Electrodynamics.* 3rd Ed. Hoboken, NJ: John Wiley and Sons, 1999.

Joo, S., D. R. Worsnop, C. E. Kolb, S. K. Kim, and D. R. Herschbach, 1999. "Observation of the A2 Sigma + <--X2Pi Electronic Transition of NaO." *Journal of Physical Chemistry A* 103, 3193-3199.

Krane, Kenneth S. *Introductory Nuclear Physics.* Hoboken, NJ: John Wiley and Sons, 1988.

Li, T. "Sodium Lidar Observed Variability in Mesopause Region Temperature and Horizontal Wind: Planetary Wave Influence and Tidal-Gravity Wave Interactions." *PhD Dissertation, Colorado State University.* 2005.

Lindzen, R. S, 1981. "Turbulance and stress owing to gravity wave and tidal breakdown." *J. Geophys. Res.* 86, 9707-9714.

Macaluso, D., and O. M. Corbino, 1898. "Sopra una nuova azione che la luce subisce attraversando alcuni vapori metallici in un campo magnetico." *Il Nuovo Cimento* IV, 257-258.

McIlrath, T. J., R. Hudson, R. Aikin, and T. D. Wilkerson, 1979. "Two-photon lidar technique for rempte sensing of atomic oxygen." *Applied Optics* 18, 316-319.

Melo, Stella M. L., Ian C. McDade, and Hisao Takahasi, 2001. "Atomic oxygen density profiles from ground-based nightglow measurements at 23 S." *J. Geophys. Res.* 106, 15377-15384.

Menders, J. K., S. Benson, H. Bloom, C. S. Liu, and Eric Korevaar, 1991. "Ultranarrow line filtering using a Cs Faraday filter at 852 nm." *Optics Letters* 16, 846-848.

Offermann, D., V. Friedrich, P. Ross, and U. von Zahn, 1981. "Neutral gas composition measurements between 80 and 120 km." *Planetary and Space Science* 29, 747-764.

Ohman, Yngve, 1956. "On some new auxiliary instruments in astrophysical research." *Stockholms Observatoriums Annaler* 19, no. 4, 1-14.

Palo, S. "Meteors, Meteor Radar and Mesospheric Winds." *CEDAR Workshop.* Santa Fe, NM, 2007.

Plane, J. M. C. *Personal communication* (February 3, 2010).

Plane, J. M. C, 2003. "Atmospheric Chemistry of Metoric Metals." *Chem. Rev.* 103, 4963-4984.

Plane, J. M. C. *Personal communication* (May 10, 2010).

Press, William H., Saul A. Teukolsky, William T. Vetterling, and Brian P. Flannery. *Numerical Recipes in C: The Art of Scientific Computing.* 2nd. Cambridge: Cambridge University Press, 1992.

Ralchenko, Yu., A.E. Kramida, J. Reader, and NIST ASD Team. *NIST Atomic Spectra Database.* National Institute of Technology, Gaithersburg, MD. 2008. http://physics.nist.gov/asd3.

Rayleigh, Lord, 1924. "The light of the Night Sky: Its intensity variations when analysed by colour filters." *Proceedings of the Royal Society, A* 106, 117-137.

Roach, F. E., and H. B. Pettit, 1951. "The annual variation of sodium D in the nightglow." *Ann. Astrophys.* 14, 392-398.

Roberts, J. L. *Personal Communication* (June 16, 2010).

Rogers, W. A., and M. A. Biondi, 1964. "Studies of the Mechanism of Electron-Ion Recombination. II." *Physical Review* 134, A1215.

Sahr, J. D. "Introduction to Ionospheric Radar Remote Sensing." *CEDAR Workshop.* Santa Fe, NM, 2006.

Sakurai, J. J. *Modern Quantum Mechanics.* Revised Ed. Addison-Wesley, 1994.

Sargoytchev, Stoyan I., Stephen Brown, Brian H. Solheim, Young-Min Cho, Gordon G. Shepherd, and Maria Jose Lopez-Gonzalez, 2004. "Spectral airglow temperature imager (SATI): a ground-based instrument for the monitoring of mesosphere temperature." *Applied Optics* 43, 5712-5721.

Schreier, F, 1992. "The Voigt and Compex Error Function: A Comparison of Computational Methods." *J. Quant. Spectrosc. Radiat. Transfer* 48, 743-762.

She, C. Y., and J. R. Yu, 1994. "Simultaneous three-frequency Na lidar measurements of radial wind and temperature in the mesopause region." *Geophys. Res. Lett.* 21, 1771-1774.

She, C. Y., et al, 2006. "Simultaneous observation of sodium atoms, NLC and PMSE in the summer mesopause region above ALOMAR, Norway (69N, 12E)." *J Atmos Solar-Terr Phys* 68, 93-101.

She, C. Y., J. R. Yu, H. Latifi, and R. E. Bills, 1992. "High-spectral-resolution fluorescence light detection and ranging for mesospheric sodium temperature measurements." *Applied Optics* 31, 2095-2106.

She, C.-Y., and J. R. Yu, 1995. "Doppler-free saturation fluorescence spectroscopy of Na atoms for atmospheric application." *Applied Optics* 34, 1063-1075.

She, C.-Y., et al, 2004. "Tidal perturbations and variability in the mesopause region over Fort Collins (41N, 105W): Continuous multi-day temperature and wind lidar observations." *Geophys. Res. Lett.* 34, L24111.

Shi, X., D. R. Herschbach, D. R. Worsnop, and C. E. Kolb, 1993. "Molecular Beam Chemistry: Magnetic Deflection Analysis of Monoxide Electronic States from Alkali-Metal Atom + Ozone Reactions." *J. Phys. Chem* 97, 2113-2122.
Sipler, Dwight P., and Manfred A. Biondi, 1978. "Interferometric Studies of the Twighlight and Nightglow Sodium D-Line Profiles." *Planet. Space Sci.* 26, 65-73.

Slanger, T. G., and J. M. C. Plane, 2005. "The D2/D1 sodium nightglow intensity ratio as a mesospheric probe." *Remote Sensing of the Clouds and the Atmosphere X, Proc. SPIE.*, 59790Y-1.

Slanger, T. G., et al, 2005. "Variability of the mesospheric nightglow sodium D2/D1 ratio." *J. Geophys. Res.* 110, D23302.

Slipher, V. M, 1929. "Emmissions in the spectrum of the light of the night sky." *Pulbl. Ast. Soc. Pacif.* 41, 262-263.

Smith, Anne K, 2004. "Physics and Chemistry of the mesopause region." *Journal of Atmospheric and Solar-Terrestrial Physics* 66, 839-857.

Stone, N. J, 2005. "Table of nuclear magnetic dipole and electric quadrupole moments." *Atomic Data and Nuclear Data Tables* 90, 75-176.

Taylor, John R. *An Introduction to Error Analysis.* Sausalito, CA: University Science Books, 1997.

Thomas, R. J, 1990. "Atomic Hydrogen and Atomic Oxygen Density in the Mesopause Region: Global and Seasonal Variations Deduced From Solar Mesosphere Explorer Near-Infrared Emissions." *Journal of Geophysical Research* 95, 16457-16476.

Van Baak, D. A, 1996. "Resonant Faraday rotation as a probe of atomic dispersion." *American Journal of Physics* 64, 724-735.

Van Wijngaarden, W. A., and J. Li, 1994. "Measurement of hyperfine structure of sodium 3P1/2,3/2 states using optical spectroscopy." *Zeitschrift fur Physik D* 32, 67-71.

Wiens, R. H., and G. Weill, 1973. "Diurnal, annual and solar cycle variations of hydroxyl and sodium nightglow intensities in the Europe-Africa sector." *Planet. Space Sci.* 21, 1011-1027.

Wiese, W. L., M. W. Smith, and B. M. Miles. "Atomic Transition Probabilities: Volume II, Sodium Through Calcium." Nat. Stand. Ref. Data Ser., Nat. Bur. Stand., U.S.

Williams, B. P. "Observations of the Atmospheric Tide, Mean Wind, and Sodium Nightglow near the Mesopause with the Magneto-Optic Doppler Analyzer." Boulder, CO: University of Colorado, 1997.

Williams, B. P., and S. Tomczyk, 1996. "Magneto-optic Doppler analyzer: a new instrument to measure mesopause winds." *Applied Optics* 35, 6494-6503.

Yee, J. H., and W. Skinner. "Spatial and Temporal Variations of the Atomic Oxygen in the Mesosphere and Lower Thermosphere." *TIMED Science Meeting.* Boulder, CO, 2007.

Yee, J. H., G. Crowley, R. G. Roble, W. R. Skinner, M. D. Burrage, and P. B. Hays, 1997. "Global simulations and observations of O(1S), O2(1Sigma) and OH mesospheric nightglow emissions." *Jornal of Geophysical Research* 102, 19949-19968.

Yeh, Pohci, 1982. "Dispersive magnetooptic filters." *Applied Optics* 21, 2069-2075.

Yin, B., and T. M. Shay, 1991. "Theoretical Model for a Faraday anomalous dispersion optical filter." *Optics Letters* 16, 1617-1619.

Yntema, L, 1909. "On the brightness of the sky and total amount of starlight." *Publications of the Astronomical Laboratory at Groningen* 22 , 1-55.

Yuan, T, C. Y. She, H. Schmidt, D. A. Krueger, and S. Reising, 2008. "Seasonal variations of semidiurnal tidal-period perturbations in mesopause region temperature, zonal and meridional winds above Fort Collins, CO (40.6N, 105W)." *J. Geophys. Res* 113, D20103.

Yuan, T, et al, 2006. "Seasonal variation of diurnal perturbations in mesopause region temperature, zonal, and meridional winds above Fort Collins, Colorado (40.6N, 105W)." *J. Geophys. Res.* 111, D06103.

Yuan, Tao. "Seasonal variations of diurnal and semidiurnal tidal-period perturbations in mesopause region temperature and zonal and meridional winds above Ft. Collins, CO (40N, 105W) based on Na-Lidar observation over full diurnal cycles." *Dissertation.* 2004.

Yuan, Tao, et al, 2009. "Wind-Bias correction method for narrowband sodium Doppler lidars using iodine absorption spectroscopy." *Applied Optics* 48, 1-6.

Yue, Jia. "Effects of background wind and temperature on bores, strong wind shears, and concentric gravity waves in the mesopause region." *PhD Dissertation.* Colorado State University, 2009.

Yue, Jia. *Personal Communication* (May 19, 2010).

Zhang, Yundong, Xiaoling Jia, Zuguang Ma, and Qi Wang, 2001:. "Optical Filtering Characteristic of Potassium Faraday Optical Filter." *IEEE Journal of Quantum Electronics,* 37, 372-375.

APPENDIX A

DETAILS OF THE QUANTUM-MECHANICAL CALCULATION OF ELECTRONIC TRANSITIONS

A.1 Derivation of transition matrix elements and relation to known parameters

The starting point is the $|m_J\, m_I\rangle$ base states of section 4.2.2. As an example, the states $^2S_{1/2}$ and $^2P_{1/2}$ have J=1/2 (I=3/2 for Na and K). The set of 8 $|m_I\, m_J\rangle$ eigenstates is:

$$\overset{1}{\left|\frac{3}{2}\frac{1}{2}\right\rangle}\overset{2}{\left|\frac{3}{2}\frac{-1}{2}\right\rangle}\overset{3}{\left|\frac{1}{2}\frac{1}{2}\right\rangle}\overset{4}{\left|\frac{1}{2}\frac{-1}{2}\right\rangle}\overset{5}{\left|\frac{-1}{2}\frac{1}{2}\right\rangle}\overset{6}{\left|\frac{-1}{2}\frac{-1}{2}\right\rangle}\overset{7}{\left|\frac{-3}{2}\frac{1}{2}\right\rangle}\overset{8}{\left|\frac{-3}{2}\frac{-1}{2}\right\rangle}, \tag{A.1}$$

and the numbers above the kets will be used as a simpler notation. We now proceed to solve the eigenvalue problem for the total Hamiltonian.

Raising and lowering operators, $I_\pm$ and $J_\pm$, allow H_I to be written as a sum of three terms, according to the power of raising or lowering operators:

$$\begin{aligned}
H_{\pm 0} &= g_J \mu_B B_0 J_z - g_I \mu_I B_0 I_z + A_J I_z J_z + \\
&\quad \frac{B_J}{2I(I-1)J(2J-1)}\left[3\left(I_z J_z I_z J_z + \frac{1}{4}\left(I_+ J_- I_- J_+ + I_- J_+ I_+ J_-\right)\right) + \frac{3}{2} I_z J_z - I(I+1)J(J+1)\right], \\
H_{\pm 1} &= \frac{A_J}{2}\left(I_+ J_- + I_- J_+\right) + \\
&\quad \frac{B_J}{2I(I-1)J(2J-1)}\frac{3}{2}\left[I_z J_z\left(I_+ J_- + I_- J_+\right) + \left(I_+ J_- + I_- J_+\right) I_z J_z + \frac{1}{2}\left(I_+ J_- + I_- J_+\right)\right], \\
H_{\pm 2} &= \frac{B_J}{2I(I-1)J(2J-1)}\frac{3}{4}\left(I_+ J_- I_+ J_- + I_- J_+ I_- J_+\right).
\end{aligned} \tag{A.2}$$

Continuing with the $^2S_{1/2}$ (or $^2P_{1/2}$) example, since there are 8 states, the Hamiltonian will be an 8x8 matrix in block-diagonal form, with the subscripts using the notation of Eq. (A.1):

$$H_{HFS} = \begin{pmatrix} H_{1,1} & & & & & & & 0 \\ & H_{2,2} & H_{2,3} & & & & & \\ & H_{3,2} & H_{3,3} & & & & & \\ & & & H_{4,4} & H_{4,5} & & & \\ & & & H_{5,4} & H_{5,5} & & & \\ & & & & & H_{6,6} & H_{6,7} & \\ & & & & & H_{7,6} & H_{7,7} & \\ 0 & & & & & & & H_{8.8} \end{pmatrix}. \tag{A.3}$$

The values for the Hamiltonian matrix elements for these states of Na are listed here:

$$\begin{aligned} H_{1,1} &= \frac{3}{4}A_J + \frac{1}{2}g_J\mu_B B_0 - \frac{3}{2}g_I\mu_n B_0, \\ H_{2,2} &= -\frac{3}{4}A_J - \frac{1}{2}g_J\mu_B B_0 - \frac{3}{2}g_I\mu_n B_0, \\ H_{3,3} &= \frac{1}{4}A_J + \frac{1}{2}g_J\mu_B B_0 - \frac{1}{2}g_I\mu_n B_0, \\ H_{4,4} &= -\frac{1}{4}A_J - \frac{1}{2}g_J\mu_B B_0 - \frac{1}{2}g_I\mu_n B_0, \\ H_{5,5} &= -\frac{1}{4}A_J + \frac{1}{2}g_J\mu_B B_0 + \frac{1}{2}g_I\mu_n B\ , \\ H_{6,6} &= \frac{1}{4}A_J - \frac{1}{2}g_J\mu_B B_0 + \frac{1}{2}g_I\mu_n B_0, \\ H_{7,7} &= -\frac{3}{4}A_J + \frac{1}{2}g_J\mu_B B_0 + \frac{3}{2}g_I\mu_n B_0, \\ H_{8,8} &= \frac{3}{4}A_J - \frac{1}{2}g_J\mu_B B_0 + \frac{3}{2}g_I\mu_n B_0, \\ H_{2,3} &= H_{3,2} = H_{6,7} = H_{7,6} = \frac{\sqrt{3}}{2}A_J, \\ H_{4,5} &= H_{5,4} = A_J. \end{aligned} \tag{A.4}$$

The blocks in the Hamiltonian matrices, at most 2x2, can be solved independently and easily, yielding energy eigenvalues E and eigenstates with their associated coefficients. As an example, the solution of one block in the matrix is presented here:

$$\begin{pmatrix} H_{2,2} & H_{2,3} \\ H_{3,2} & H_{3,3} \end{pmatrix} \begin{pmatrix} a_{-2,3} \\ a_{+2,3} \end{pmatrix} = \lambda \begin{pmatrix} a \\ b \end{pmatrix}, \tag{A.5}$$

where λ represents the eigenvalues, and $\begin{pmatrix} a_{-2,3} \\ a_{+2,3} \end{pmatrix}$ are the eigenstate coefficients. Solving Eq. (A.5) for the eigenvalues and eigenstate coefficients yields

$$\begin{aligned} \lambda &= \frac{1}{2}\left(H_{2,2} + H_{3,3}\right) \pm \sqrt{\left(\frac{H_{2,2} - H_{3,3}}{2}\right)^2 + H_{2,3}H_{3,2}}, \\ a_{+2,3} &= \frac{H_{2,3}}{\sqrt{H_{2,3}^2 + \left(H_{2,2} - \lambda\right)^2}}, \; a_{-2,3} = \frac{\lambda - H_{2,2}}{H_{2,3}} a_{+2,3} = \frac{\lambda - H_{2,2}}{\sqrt{H_{2,3}^2 + \left(H_{2,2} - \lambda\right)^2}}. \end{aligned} \tag{A.6}$$

In terms of the $|m_I m_J\rangle$ basis kets, the eigenstates become:

$$\begin{aligned} \left|2\right\rangle &= a_{-2}\left|\frac{3}{2},-\frac{1}{2}\right\rangle + a_{+2}\left|\frac{1}{2},\frac{1}{2}\right\rangle, \\ \left|3\right\rangle &= a_{-3}\left|\frac{3}{2},-\frac{1}{2}\right\rangle + a_{+3}\left|\frac{1}{2},\frac{1}{2}\right\rangle, \end{aligned} \tag{A.7}$$

with

$$\begin{aligned} a_{-2} &= \frac{H_{2,3}}{\sqrt{H_{2,3}^2 + \left(H_{2,2} - \Delta E_2\right)^2}}, \; a_{+2} = \frac{\Delta E_2 - H_{2,2}}{\sqrt{H_{2,3}^2 + \left(H_{2,2} - \Delta E_2\right)^2}}, \\ a_{-3} &= \frac{H_{2,3}}{\sqrt{H_{2,3}^2 + \left(H_{2,2} - \Delta E_3\right)^2}}, \; a_{+3} = \frac{\Delta E_3 - H_{2,2}}{\sqrt{H_{2,3}^2 + \left(H_{2,2} - \Delta E_3\right)^2}}, \end{aligned} \tag{A.8}$$

where

$$\Delta E_2 = \frac{1}{2}\left(H_{2,2} + H_{3,3}\right) + \sqrt{\left(\frac{H_{2,2} - H_{3,3}}{2}\right)^2 + H_{2,3}H_{3,2}}\,,$$
$$\Delta E_3 = \frac{1}{2}\left(H_{2,2} + H_{3,3}\right) - \sqrt{\left(\frac{H_{2,2} - H_{3,3}}{2}\right)^2 + H_{2,3}H_{3,2}}\,. \tag{A.9}$$

The other blocks in Eq. (A.3) can be solved similarly.

For the $^2P_{3/2}$ state, there are 16 $|m_I m_J>$ states, yielding the following block diagonal matrix:

$$H_{HFS} = \begin{pmatrix}
H_{1,1} & & & & & & & & & & & & & & & \\
 & H_{2,2} & H_{2,3} & & & & & & & & & & & & & \\
 & H_{3,2} & H_{3,3} & & & & & & & & & & & & & \\
 & & & H_{4,4} & H_{4,5} & H_{4,6} & & & & & & & & & & \\
 & & & H_{4,5} & H_{5,5} & H_{5,6} & & & & & & & & & & \\
 & & & H_{6,4} & H_{6,5} & H_{6,6} & & & & & & & & & & \\
 & & & & & & H_{7,7} & H_{7,8} & H_{7,9} & H_{7,10} & & & & & & \\
 & & & & & & H_{8,7} & H_{8,8} & H_{8,9} & H_{8,10} & & & & & & \\
 & & & & & & H_{9,7} & H_{9,8} & H_{9,9} & H_{9,10} & & & & & & \\
 & & & & & & H_{10,7} & H_{10,8} & H_{10,9} & H_{10,10} & & & & & & \\
 & & & & & & & & & & H_{11,11} & H_{11,12} & H_{11,13} & & & \\
 & & & & & & & & & & H_{12,11} & H_{12,12} & H_{12,13} & & & \\
 & & & & & & & & & & H_{13,11} & H_{13,12} & H_{13,13} & & & \\
 & & & & & & & & & & & & & H_{14,14} & H_{14,15} & \\
 & & & & & & & & & & & & & H_{15,14} & H_{15,15} & \\
 & & & & & & & & & & & & & & & H_{16,16}
\end{pmatrix}. \tag{A.10}$$

The values of these matrix elements for sodium are listed in Table A.1. Each of the blocks in Eq. (A.10) can be solved to determine eigenstate coefficients, and their eigenvalues ΔE.

Table A.1: Matrix elements for the $^2P_{3/2}$ state of Na.

Eigenstate $\|m_I,m_J\rangle$	Diagonal matrix elements for $^2P_{3/2}$ including nuclear interaction with B-field	Off-diagonal matrix elements for $^2P_{3/2}$
$\|3/2,3/2\rangle 1$	$H11=2\mu_B B_0-(3/2)g_I\mu_N B_0+(9/4)A_J+(1/4)B_J$	H23=H32=
$\|3/2,1/2\rangle 2$	$H22=(2/3)\mu_B B_0-(3/2)g_I\mu_N B_0+(3/4)A_J-(1/4)B_J$	H45=H54=
$\|1/2,3/2\rangle 3$	$H33=2\mu_B B_0-(1/2)g_I\mu_N B_0+(3/4)A_J-(1/4)B_J$	$H46=H64=B_J/2$
$\|3/2,-1/2\rangle 4$	$H44=-(2/3)\mu_B B_0-(3/2)g_I\mu_N B_0-(3/4)A_J-(1/4)B_J$	H56=H65=
$\|1/2,1/2\rangle 5$	$H55=(2/3)\mu_B B_0-(1/2)g_I\mu_N B_0+(1/4)A_J+(1/4)B_J$	H78=H87=
$\|-1/2,3/2\rangle 6$	$H66=2\mu_B B_0+(1/2)g_I\mu_N B_0-(3/4)A_J-(1/4)B_J$	$H79=H97==B_J/2$
$\|3/2,-3/2\rangle 7$	$H77=-2\mu_B B_0-(3/2)g_I\mu_N B_0-(9/4)A_J+(1/4)B_J$	H710=H107=0
$\|1/2,-1/2\rangle 8$	$H88=-(2/3)\mu_B B_0-(1/2)g_I\mu_N B_0-(1/4)A_J+(1/4)B_J$	$H89=H98=2A_J$
$\|-1/2,1/2\rangle 9$	$H99=(2/3)\mu_B B_0+(1/2)g_I\mu_N B_0-(1/4)A_J+(1/4)B_J$	$H810=H108=B_J/2$
$\|-3/2,3/2\rangle 10$	$H1010=2\mu_B B_0+(3/2)g_I\mu_N B_0-(9/4)A_J+(1/4)B_J$	H910=H109=
$\|1/2,-3/2\rangle 11$	$H1111=-2\mu_B B_0-(1/2)g_I\mu_N B_0-(3/4)A_J-(1/4)B_J$	H1112=H1211=
$\|-1/2,-1/2\rangle 12$	$H1212=-(2/3)\mu_B B_0+(1/2)g_I\mu_N B_0+(1/4)A_J+(1/4)B_J$	$H1113=H1311=B_J/2$
$\|-3/2,1/2\rangle 13$	$H1313=(2/3)\mu_B B_0+(3/2)g_I\mu_N B_0-(3/4)A_J-(1/4)B_J$	H1213=H1312=
$\|-1/2,-3/2\rangle 14$	$H1414=-2\mu_B B+(1/2)g_I\mu_N B+(3/4)A_J-(1/4)B_J$	H1415=H1514=
$\|-3/2,-1/2\rangle 15$	$H1515=-(2/3)\mu_B B_0+(3/2)g_I\mu_N B_0+(3/4)A_J-(1/4)B_J$	
$\|-3/2,-3/2\rangle 16$	$H1616=-2\mu_B B_0+(3/2)g_I\mu_N B_0+(9/4)A_J+(1/4)B_J$	

With the energy levels and eigenstate coefficients determined, we can calculate the transition frequencies and probabilities of the allowed transitions. The transitions of interest are electric dipole with selection rules $\Delta m_I = 0$ and $\Delta m_J = \pm 1$ for absorption or emission of a circularly polarized photon in the filter. The transition probability is given by the square of the transition matrix element $\left|\langle m_I' m_J' | p_\pm | m_I m_J \rangle\right|^2$ where primed and unprimed represent excited and ground states. This may be written as product of 3-*j* symbols and the reduced matrix element <*I'J'*||*p*||*IJ*> (Edmonds 1957). This can be further reduced to the reduced matrix element between states in the "basic" model of the atom using 6-*j* symbols (Edmonds 1957). This basic model reduced matrix element,

$\left|\langle \ell' \| p \| \ell \rangle\right|^2$, is known as the line strength, S_0, of the transition; it is the total intensity of the line and is experimentally determined and tabulated in an NBS publication (Wiese, et al.) and reproduced in Table 4.4. The total transition probability of each Zeeman transition is given by:

$$\begin{aligned}\left|\langle m_I' m_J' \left| p_\pm \right| m_I m_J \rangle\right|^2 &= |a|^2 |b|^2 (2J'+1)(2J+1) \begin{pmatrix} J' & 1 & J \\ -m_J' & \pm 1 & m_J \end{pmatrix}^2 \begin{bmatrix} \ell' & J' & S \\ J & \ell & 1 \end{bmatrix}^2 S_0 \\ &= F_1 F_2 (2J'+1)(2J+1) \begin{bmatrix} \ell' & J' & S \\ J & \ell & 1 \end{bmatrix}^2 S_0,\end{aligned}$$

(A.11)

where a and b are the expansion coefficients of the appropriate eigenstate for the ground and excited states as defined in eq (A.8). The 3-j symbol is $\begin{pmatrix} J' & 1 & J \\ -m_J' & \pm 1 & m_J \end{pmatrix}$ and $\begin{bmatrix} \ell' & J' & S \\ J & \ell & 1 \end{bmatrix}$ is the 6-j symbol; their values are from Tables 2 and 5 in Edmonds (1957). F_1 and F_2 are defined as:

$$\begin{aligned} F_1 &= |a|^2 |b|^2, \\ F_2 &= \begin{pmatrix} J' & 1 & J \\ -m_J' & \pm 1 & m_J \end{pmatrix}^2. \end{aligned}$$

(A.12)

These probabilities are tabulated for the Na and K Zeeman-split hyperfine transitions for each circular polarization component of the $^2P_{1/2} \rightarrow {}^2S_{1/2}$ (D_1) and $^2P_{3/2} \rightarrow {}^2S_{1/2}$ (D_2) transitions, as discussed below. Note that with the choice of the $|I, J, m_I, m_J\rangle$ base states, the electronic dipole moments of the transitions are independent from the nuclear spin of the system.

A.2 D_1 transition probabilities of allowed transitions

For this transition J'=1/2, J=1/2, ℓ'=1 (P-state), ℓ=0 (S-state), and S=1/2. By inserting these values into Eq. (A5), the transition probability is found to be:

$$\left|\left\langle m_I' m_J' \left| p_\pm \right| m_I m_J \right\rangle\right|^2 = F_1 F_2 \frac{2}{3} S_0. \tag{A.13}$$

with F_2=1/3 from Eq. (A.12). Table A.2 shows the values of F_1 of the allowed transitions as well as for which $\sigma_\pm$ circular polarization. The subscripts on the a and b coefficients represent which state and m_J value they correspond to.

Table A.2: D_1 allowed transition F_1 value and polarization.

$^2S_{1/2}$	$^2P_{1/2}$	F_1	Polarization
$\|3/2,1/2\rangle$	$b_{2-}\|3/2,-1/2\rangle+b_{2+}\|1/2,1/2\rangle$	$\|b_{2-}\|^2$	$\hat{=}$
	$b_{3-}\|3/2,-1/2\rangle+b_{3+}\|1/2,1/2\rangle$	$\|b_{3-}\|^2$	$\hat{=}$
$a_{2-}\|3/2,-1/2\rangle+a_{2+}\|1/2,1/2\rangle$	$\|3/2,1/2\rangle$	$\|a_{2-}\|^2$	$\hat{+}$
	$b_{4-}\|1/2,-1/2\rangle+b_{4+}\|-1/2,1/2\rangle$	$\|a_{2+}\|^2\|b_{4-}\|^2$	$\hat{=}$
	$b_{5-}\|1/2,-1/2\rangle+b_{5+}\|-1/2,1/2\rangle$	$\|a_{2+}\|^2\|b_{5-}\|^2$	$\hat{=}$
$a_{3-}\|3/2,-1/2\rangle+a_{3+}\|1/2,1/2\rangle$	$\|3/2,1/2\rangle$	$\|a_{3-}\|^2$	$\hat{+}$
	$b_{4-}\|1/2,-1/2\rangle+b_{4+}\|-1/2,1/2\rangle$	$\|a_{3+}\|^2\|b_{4-}\|^2$	$\hat{=}$
	$b_{5-}\|1/2,-1/2\rangle+b_{5+}\|-1/2,1/2\rangle$	$\|a_{3+}\|^2\|b_{5-}\|^2$	$\hat{=}$
$a_{4-}\|1/2,-1/2\rangle+a_{4+}\|-1/2,1/2\rangle$	$b_{2-}\|3/2,-1/2\rangle+b_{2+}\|1/2,1/2\rangle$	$\|a_{4-}\|^2\|b_{2+}\|^2$	$\hat{+}$
	$b_{3-}\|3/2,-1/2\rangle+b_{3+}\|1/2,1/2\rangle$	$\|a_{4-}\|^2\|b_{3+}\|^2$	$\hat{+}$
	$b_{6-}\|-1/2,-1/2\rangle+b_{6+}\|-3/2,1/2\rangle$	$\|a_{4+}\|^2\|b_{6-}\|^2$	$\hat{=}$
	$b_{7-}\|-1/2,-1/2\rangle+b_{7+}\|-3/2,1/2\rangle$	$\|a_{4+}\|^2\|b_{7-}\|^2$	$\hat{=}$
$a_{5-}\|1/2,-1/2\rangle+a_{5+}\|-1/2,1/2\rangle$	$b_{2-}\|3/2,-1/2\rangle+b_{2+}\|1/2,1/2\rangle$	$\|a_{5-}\|^2\|b_{2+}\|^2$	$\hat{+}$
	$b_{3-}\|3/2,-1/2\rangle+b_{3+}\|1/2,1/2\rangle$	$\|a_{5-}\|^2\|b_{3+}\|^2$	$\hat{+}$
	$b_{6-}\|-1/2,-1/2\rangle+b_{6+}\|-3/2,1/2\rangle$	$\|a_{5+}\|^2\|b_{6-}\|^2$	$\hat{=}$
	$b_{7-}\|-1/2,-1/2\rangle+b_{7+}\|-3/2,1/2\rangle$	$\|a_{5+}\|^2\|b_{7-}\|^2$	$\hat{=}$
$a_{6-}\|-1/2,-1/2\rangle+a_{6+}\|-3/2,1/2\rangle$	$b_{4-}\|1/2,-1/2\rangle+b_{4+}\|-1/2,1/2\rangle$	$\|a_{6-}\|^2\|b_{4+}\|^2$	$\hat{+}$
	$b_{5-}\|1/2,-1/2\rangle+b_{5+}\|-1/2,1/2\rangle$	$\|a_{6-}\|^2\|b_{5+}\|^2$	$\hat{+}$
	$\|-3/2,-1/2\rangle$	$\|a_{6+}\|^2$	$\hat{=}$
$a_{7-}\|-1/2,-1/2\rangle+a_{7+}\|-3/2,1/2\rangle$	$b_{4-}\|1/2,-1/2\rangle+b_{4+}\|-1/2,1/2\rangle$	$\|a_{7-}\|^2\|b_{4+}\|^2$	$\hat{+}$
	$b_{5-}\|1/2,-1/2\rangle+b_{5+}\|-1/2,1/2\rangle$	$\|a_{7-}\|^2\|b_{5+}\|^2$	$\hat{+}$
	$\|-3/2,-1/2\rangle$	$\|a_{7+}\|^2$	$\hat{=}$
$\|-3/2,-1/2\rangle$	$b_{6-}\|-1/2,-1/2\rangle+b_{6+}\|-3/2,1/2\rangle$	$\|b_{6+}\|^2$	$\hat{+}$
	$b_{7-}\|-1/2,-1/2\rangle+b_{7+}\|-3/2,1/2\rangle$	$\|b_{7+}\|^2$	$\hat{+}$

A.3 D_2 transition probabilities of allowed transitions

For this transition J'=3/2, J=1/2, ℓ'=1 (P-state), ℓ=0 (S-state), and S=1/2. By inserting these values into Eq. (A5), the transition probability is found to be:

$$\left|\left\langle m_I' m_J' \left| p_{\pm} \right| m_I m_J \right\rangle\right|^2 = F_1 F_2 \frac{4}{3} S_0. \tag{A.14}$$

where F_2= 1/12 or 1/4, depending on whether $|m_J'|$ = 1/2, or 3/2, respectively, as indicated in Table A.3. Table A.3 shows the values of F_1F_2 of the allowed transitions as well as the $\sigma_{\pm}$ circular polarization. The subscripts on the a and b coefficients represent which state and m_J value they correspond to.

Table A.3: D_2 allowed transitions F_1 and F_2 values and polarization

$^2S_{1/2}$	$^2P_{3/2}$	F_1F_2	Pol.
$\|3/2,1/2\rangle$	$\|3/2,3/2\rangle$	¼	$\hat{+}$
	$b_{4,-.5}\|3/2,-1/2\rangle+b_{4,.5}\|1/2,1/2\rangle+b_{4,1.5}\|-1/2,3/2\rangle$	$\|b_{4,-.5}\|^2/12$	$\hat{=}$
	$b_{5,-.5}\|3/2,-1/2\rangle+ b_{5,.5}\|1/2,1/2\rangle+ b_{5,1.5}\|-1/2,3/2\rangle$	$\|b_{5,-.5}\|^2/12$	$\hat{=}$
	$b_{6,-.5}\|3/2,-1/2\rangle+ b_{6,.5}\|1/2,1/2\rangle+ b_{6,1.5}\|-1/2,3/2\rangle$	$\|b_{6,-.5}\|^2/12$	$\hat{=}$
$a_{2-}\|3/2,-1/2\rangle+a_{2+}\|1/2,1/2\rangle$	$b_{2,.5}\|3/2,1/2\rangle + b_{2,1.5}\|1/2,3/2\rangle$	$\|a_{2-}\|^2\|b_{2,.5}\|^2/12+\|a_{2+}\|^2\|b_{2,1.5}\|^2/4$	$\hat{+}$
	$b_{3,.5}\|3/2,1/2\rangle + b_{3,1.5}\|1/2,3/2\rangle$	$\|a_{2-}\|^2\|b_{3,.5}\|^2/12+\|a_{2+}\|^2\|b_{3,1.5}\|^2/4$	$\hat{+}$
	$b_{7,-1.5}\|3/2,-3/2\rangle+b_{7,-.5}\|1/2,-1/2\rangle+b_{7,.5}\|-1/2,1/2\rangle +b_{7,1.5}\|-3/2,3/2\rangle$	$\|a_{2-}\|^2\|b_{7,-1.5}\|^2/4+\|a_{2+}\|^2\|b_{7,-0.5}\|^2/12$	$\hat{=}$
	$b_{8,-1.5}\|3/2,-3/2\rangle+b_{8,-.5}\|1/2,-1/2\rangle+b_{8,.5}\|-1/2,1/2\rangle+ b_{8,1.5}\|-3/2,3/2\rangle$	$\|a_{2-}\|^2\|b_{8,-1.5}\|^2/4+ \|a_{2+}\|^2\|b_{8,-0.5}\|^2/12$	$\hat{=}$
	$b_{9,-1.5}\|3/2,-3/2\rangle+b_{9,-.5}\|1/2,-1/2\rangle+b_{9,.5}\|-1/2,1/2\rangle +b_{9,1.5}\|-3/2,3/2\rangle$	$\|a_{2-}\|^2\|b_{9,-1.5}\|^2/4+ \|a_{2+}\|^2\|b_{9,-0.5}\|^2/12$	$\hat{=}$
	$b_{10,-1.5}\|3/2,-3/2\rangle+b_{10,-.5}\|1/2,-1/2\rangle+b_{10,.5}\|-1/2,1/2\rangle+b_{10,1.5}\|-3/2,3/2\rangle$	$\|a_{2-}\|^2\|b_{10,-1.5}\|^2/4+ \|a_{2+}\|^2\|b_{10,-0.5}\|^2/12$	$\hat{=}$
$a_{3-}\|3/2,-1/2\rangle+a_{3+}\|1/2,1/2\rangle$	$b_{2,.5}\|3/2,1/2\rangle + b_{2,1.5}\|1/2,3/2\rangle$	$\|a_{3-}\|^2\|b_{2,0.5}\|^2/12 +\|a_{3+}\|^2\|b_{2,1.5}\|^2/4$	$\hat{+}$
	$b_{3,.5}\|3/2,1/2\rangle + b_{3,1.5}\|1/2,3/2\rangle$	$\|a_{3-}\|^2\|b_{3,0.5}\|^2/12 +\|a_{3+}\|^2\|b_{3,1.5}\|^2/4$	$\hat{+}$
	$b_{7,-1.5}\|3/2,-3/2\rangle+b_{7,-.5}\|1/2,-1/2\rangle+ b_{7,.5}\|-1/2,1/2\rangle+b_{7,1.5}\|-3/2,3/2\rangle$	$\|a_{3-}\|^2\|b_{7,-1.5}\|^2/4+ \|a_{3+}\|^2\|b_{7,-0.5}\|^2/12$	$\hat{=}$
	$b_{8,-1.5}\|3/2,-3/2\rangle+ b_{8,-.5}\|1/2,-1/2\rangle+ b_{8,.5}\|-1/2,1/2\rangle+ b_{8,1.5}\|-3/2,3/2\rangle$	$\|a_{3-}\|^2\|b_{8,-1.5}\|^2/4+ \|a_{3+}\|^2\|b_{8,-0.5}\|^2/12$	$\hat{=}$
	$b_{9,-1.5}\|3/2,-3/2\rangle+ b_{9,-.5}\|1/2,-1/2\rangle+ b_{9,.5}\|-1/2,1/2\rangle+ b_{9,1.5}\|-3/2,3/2\rangle$	$\|a_{3-}\|^2\|b_{9,-1.5}\|^2/4+ \|a_{3+}\|^2\|b_{9,-0.5}\|^2/12$	$\hat{=}$
	$b_{10,-1.5}\|3/2,-3/2\rangle+ b_{10,-.5}\|1/2,-1/2\rangle+ b_{10,.5}\|-1/2,1/2\rangle+ b_{10,1.5}\|-3/2,3/2\rangle$	$\|a_{3-}\|^2\|b_{10,-1.5}\|^2/4+ \|a_{3+}\|^2\|b_{10\text{-}0.5}\|^2/12$	$\hat{=}$
$a_{4-}\|1/2,-1/2\rangle+a_{4+}\|-1/2,1/2\rangle$	$b_{4,-.5}\|3/2,-1/2\rangle+b_{4,.5}\|1/2,1/2\rangle+b_{4,1.5}\|-1/2,3/2\rangle$	$\|a_{4-}\|^2\|b_{4,0.5}\|^2/12 +\|a_{4+}\|^2\|b_{4,1.5}\|^2/4$	$\hat{+}$
	$b_{5,-.5}\|3/2,-1/2\rangle+b_{5,.5}\|1/2,1/2\rangle+b_{5,1.5}\|-1/2,3/2\rangle$	$\|a_{4-}\|^2\|b_{5,0.5}\|^2/12 +\|a_{4+}\|^2\|b_{5,1.5}\|^2/4$	$\hat{+}$
	$b_{6,-.5}\|3/2,-1/2\rangle+b_{6,.5}\|1/2,1/2\rangle+b_{6,1.5}\|-1/2,3/2\rangle$	$\|a_{4-}\|^2\|b_{6,0.5}\|^2/12+\|a_{4+}\|^2\|b_{6,1.5}\|^2/4$	$\hat{+}$
	$b_{11,-1.5}\|1/2,-3/2\rangle+b_{11,-.5}\|-1/2,-1/2\rangle+b_{11,.5}\|-3/2,1/2\rangle$	$\|a_{4-}\|^2\|b_{11,-1.5}\|^2/4+ \|a_{4+}\|^2\|b_{11,-0.5}\|^2/12$	$\hat{=}$
	$b_{12,-1.5}\|1/2,-3/2\rangle+b_{12,-.5}\|-1/2,-1/2\rangle+b_{12,.5}\|-3/2,1/2\rangle$	$\|a_{4-}\|^2\|b_{12,-1.5}\|^2/4+ \|a_{4+}\|^2\|b_{12,-0.5}\|^2/12$	$\hat{=}$
	$b_{13,-1.5}\|1/2,-3/2\rangle+b_{13,-.5}\|-1/2,-1/2\rangle+b_{13,.5}\|-3/2,1/2\rangle$	$\|a_{4-}\|^2\|b_{13,-1.5}\|^2/4+ \|a_{4+}\|^2\|b_{13,-0.5}\|^2/12$	$\hat{=}$
$a_{5-}\|1/2,-1/2\rangle+a_{5+}\|-1/2,1/2\rangle$	$b_{4,-.5}\|3/2,-1/2\rangle+b_{4,.5}\|1/2,1/2\rangle+b_{4,1.5}\|-1/2,3/2\rangle$	$\|a_{5-}\|^2\|b_{4,0.5}\|^2/12 +\|a_{5+}\|^2\|b_{4,1.5}\|^2/4$	$\hat{+}$
	$b_{5,-.5}\|3/2,-1/2\rangle+b_{5,.5}\|1/2,1/2\rangle+b_{5,1.5}\|-1/2,3/2\rangle$	$\|a_{5-}\|^2\|b_{5,0.5}\|^2/12 +\|a_{5+}\|^2\|b_{5,1.5}\|^2/4$	$\hat{+}$
	$b_{6,-.5}\|3/2,-1/2\rangle+b_{6,.5}\|1/2,1/2\rangle+b_{6,1.5}\|-1/2,3/2\rangle$	$\|a_{5-}\|^2\|b_{6,0.5}\|^2/12 +\|a_{5+}\|^2\|b_{6,1.5}\|^2/4$	$\hat{+}$
	$b_{11,-1.5}\|1/2,-3/2\rangle+b_{11,-.5}\|-1/2,-1/2\rangle+b_{11,.5}\|-3/2,1/2\rangle$	$\|a_{5-}\|^2\|b_{11,-1.5}\|^2/4+ \|a_{5+}\|^2\|b_{11,-0.5}\|^2/12$	$\hat{=}$
	$b_{12,-1.5}\|1/2,-3/2\rangle+b_{12,-.5}\|-1/2,-1/2\rangle+b_{12,.5}\|-3/2,1/2\rangle$	$\|a_{5-}\|^2\|b_{12,-1.5}\|^2/4+ \|a_{5+}\|^2\|b_{12,-0.5}\|^2/12$	$\hat{=}$
	$b_{13,-1.5}\|1/2,-3/2\rangle+b_{13,-.5}\|-1/2,-1/2\rangle+b_{13,.5}\|-3/2,1/2\rangle$	$\|a_{5-}\|^2\|b_{13,-1.5}\|^2/4+ \|a_{5+}\|^2\|_{b13,-0.5}\|^2/12$	$\hat{=}$

$a_{6,-}\|-1/2,-1/2\rangle+a_{6,+}\|-3/2,1/2\rangle$	$b_{7,-1.5}\|3/2,-3/2\rangle+b_{7,-.5}\|1/2,-1/2\rangle+b_{7,.5}\|-1/2,1/2\rangle+b_{7,1.5}\|-3/2,3/2\rangle$	$\|a_{6,-}\|^2\|b_{7,0.5}\|^2/12+\|a_{6,+}\|^2\|b_{7,1.5}\|^2/4$	$\hat{+}$
	$b_{8,-1.5}\|3/2,-3/2\rangle+b_{8,-.5}\|1/2,-1/2\rangle+b_{8,.5}\|-1/2,1/2\rangle+b_{8,1.5}\|-3/2,3/2\rangle$	$\|a_{6,-}\|^2\|b_{8,0.5}\|^2/12+\|a_{6,+}\|^2\|b_{8,1.5}\|^2/4$	$\hat{+}$
	$b_{9,-1.5}\|3/2,-3/2\rangle+b_{9,-.5}\|1/2,-1/2\rangle+b_{9,.5}\|-1/2,1/2\rangle+b_{9,1.5}\|-3/2,3/2\rangle$	$\|a_{6,-}\|^2\|b_{9,0.5}\|^2/12+\|a_{6,+}\|^2\|b_{9,1.5}\|^2/4$	$\hat{+}$
	$b_{10,-1.5}\|3/2,-3/2\rangle+b_{10,-.5}\|1/2,-1/2\rangle+b_{10,.5}\|-1/2,1/2\rangle+b_{10,1.5}\|-3/2,3/2\rangle$	$\|a_{6,-}\|^2\|b_{10,0.5}\|^2/12+\|a_{6,+}\|^2\|b_{10,1.5}\|^2/4$	$\hat{+}$
	$b_{14,-1.5}\|-1/2,-3/2\rangle+b_{14,-.5}\|-3/2,-1/2\rangle$	$\|a_{6,-}\|^2\|b_{14,-1.5}\|^2/4+\|a_{6,+}\|2\|b_{14,-0.5}\|^2/12$	$\hat{=}$
	$b_{15,-1.5}\|-1/2,-3/2\rangle+b_{15,-.5}\|-3/2,-1/2\rangle$	$\|a_{6,-}\|^2\|b_{15,-1.5}\|^2/4+\|a_{6,+}\|^2\|b_{15,-0.5}\|^2/12$	$\hat{=}$
$a_{7,-}\|-1/2,-1/2\rangle+a_{7,+}\|-3/2,1/2\rangle$	$b_{7,-1.5}\|3/2,-3/2\rangle+b_{7,-.5}\|1/2,-1/2\rangle+b_{7,.5}\|-1/2,1/2\rangle+b_{7,1.5}\|-3/2,3/2\rangle$	$\|a_{7,-}\|^2\|b_{7,0.5}\|^2/12+\|a_{7,+}\|^2\|b_{7,1.5}\|^2/4$	$\hat{+}$
	$b_{8,-1.5}\|3/2,-3/2\rangle+b_{8,-.5}\|1/2,-1/2\rangle+b_{8,.5}\|-1/2,1/2\rangle+b_{8,1.5}\|-3/2,3/2\rangle$	$\|a_{7,-}\|^2\|b_{8,0.5}\|^2/12+\|a_{7,+}\|^2\|b_{8,1.5}\|^2/4$	$\hat{+}$
	$b_{9,-1.5}\|3/2,-3/2\rangle+b_{9,-.5}\|1/2,-1/2\rangle+b_{9,.5}\|-1/2,1/2\rangle+b_{9,1.5}\|-3/2,3/2\rangle$	$\|a_{7,-}\|^2\|b_{9,0.5}\|^2/12+\|a_{7,+}\|^2\|b_{9,1.5}\|^2/4$	$\hat{+}$
	$b_{10,-1.5}\|3/2,-3/2\rangle+b_{10,-.5}\|1/2,-1/2\rangle+b_{10,.5}\|-1/2,1/2\rangle+b_{10,1.5}\|-3/2,3/2\rangle$	$\|a_{7,-}\|^2\|b_{10,0.5}\|^2/12+\|a_{7,+}\|^2\|b_{10,1.5}\|^2/4$	$\hat{+}$
	$b_{14,-1.5}\|-1/2,-3/2\rangle+b_{14,-.5}\|-3/2,-1/2\rangle$	$\|a_{7,-}\|^2\|b_{14,-1.5}\|^2/4+\|a_{7,+}\|^2\|b_{14,-0.5}\|^2/12$	$\hat{=}$
	$b_{15,-1.5}\|-1/2,-3/2\rangle+b_{15,-.5}\|-3/2,-1/2\rangle$	$\|a_{7,-}\|^2\|b_{15,-1.5}\|^2/4+a_{7,+}\|^2\|b_{15,-0.5}\|^2/12$	$\hat{=}$
$\|-3/2,-1/2\rangle$	$b_{11,-1.5}\|1/2,-3/2\rangle+b_{11,-.5}\|-1/2,-1/2\rangle+b_{11,.5}\|-3/2,1/2\rangle$	$\|b_{11,0.5}\|^2/12$	$\hat{+}$
	$b_{12,-1.5}\|1/2,-3/2\rangle+b_{12,-.5}\|-1/2,-1/2\rangle+b_{12,.5}\|-3/2,1/2\rangle$	$\|b_{12,0.5}\|^2/12$	$\hat{+}$
	$b_{13,-1.5}\|1/2,-3/2\rangle+b_{13,-.5}\|-1/2,-1/2\rangle+b_{13,.5}\|-3/2,1/2\rangle$	$\|b_{13,0.5}\|^2/12$	$\hat{+}$
	$\|-3/2,-3/2\rangle$	¼	$\hat{=}$

APPENDIX B

DETAILS OF THE CSU LIDAR SYSTEM

This appendix contains the specific details of the CSU sodium lidar system which are not directly applicable to the FFBS instrument. These include details of the theory of measurement and specific transmitter and receiver subsystems.

B.1 Spectral variation due to temperature and wind

As has been discussed in chapter 3, the sodium spectrum is broadened with increased temperature, due to the Doppler shift within the Maxwell-Boltzmann distribution of velocities. If the sodium atoms are in a moving atmosphere, the spectrum will also experience a Doppler shift. An example of the D_2 spectrum of sodium at various temperatures and line of sight (LOS) winds is shown in Figure B.1. One difference in the Doppler shifts due to temperature is how the broadening is manifested: the temperature broadening is due to random thermal motion, and thus the atoms could move towards or away from the observer, so the shift occurs as a broadening of the lineshape. The shift due to the mean wind is different, however; since all the atoms experience the same motion towards or away from the observer, the entire spectrum experiences a center of mass shift:

$$\Delta\nu = -V / \lambda_0 \tag{B.1}$$

where λ_0=589.158 nm.

To determine the temperature and wind from the Doppler broadening and shift, we probe the sodium spectrum at three frequencies as indicated in figure B.1: the D_{2a} peak (ν_0), one at 630 MHz above the peak (ν_+), and one at 630 MHz below the peak (ν_-). The returned lidar signals at each frequency can form two ratios, one sensitive to temperature broadening and the other sensitive to wind shift. This process will be detailed in the next section.

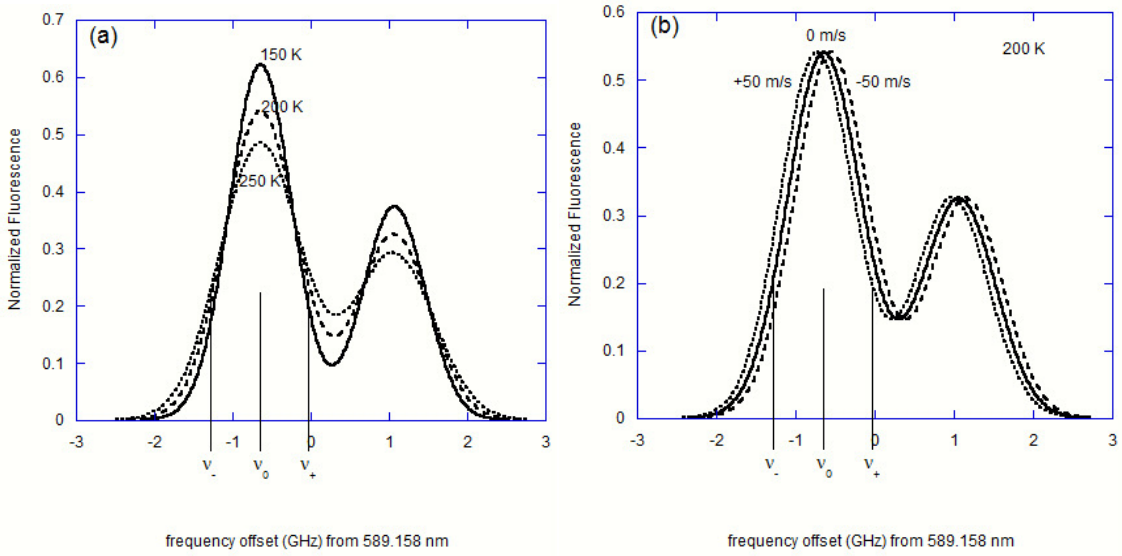

Figure B.1: The Na hyperfine D_2 spectra normalized to unit area. (a) Spectra at 150 K (solid), 200 K (dashed) and 250 K (dotted), and 0 m/s wind, showing effects of temperature broadening. (b) Spectra at 200 K, and at three radial velocities: 0 m/s (solid), +50 m/s (dotted) and -50 m/s (dashed). Also indicated are the three lidar probe frequencies.

B.2 The lidar equation

When the lidar pulses are sent into the atmosphere, the photons encounter a constant barrage of scatterers and absorbers on their way up to the mesopause. The light then interacts with the sodium atoms according to the scattering cross-section. The returned fluorescence then experiences more scattering and absorption. All of the physical processes involved between pulse transmission and detection must be accounted

for in the so-called lidar equation. For the CSU sodium fluorescence lidar, the lidar equation for the backscatter signal from Na atoms at an altitude z can be written as:

$$N(z) = \left(\eta T_A^2\right)\left(\frac{E_L}{hc/\lambda_L}\right)\left(\rho(z)\Delta z\sigma_{sb}\right)\left(\frac{A_R}{z^2}\right)\exp\left[-2\int_{z_l}^{z}\alpha(z')dz'\right] + N_B \qquad \text{(B.2)}$$

where:

$N(z)$ is the number of photons in the range bin $(z-\Delta z/2,\ z+\Delta z/2)$

η is the efficiency of the lidar system

T_A^2 is the two-way transmittance of the atmosphere between the ground and the bottom edge of the sodium layer

E_L is the energy of the emitted laser pulse

hc/λ_L is the energy of one photon at the transmitted frequency

ρ is the sodium number density

σ_{sb} is the sodium differential backscattering cross section

Δz is the range bin size

A_R is the receiving telescope area

$\alpha = 4\pi\rho\sigma_{sb}=\rho\sigma_{sa}$ is the sodium extinction coefficient and σ_{sa} is the total absorption cross section

$\exp\left[-2\int_{z_l}^{z}\alpha(z')dz'\right]$ is the two way transmittance in the sodium layer between its bottom edge and the height in question (assuming the frequency of the upward and downward photons is the same, see Acott (2009))

and

N_B is the background noise per range bin per pulse.

For daytime lidar measurements, equation B.2 needs an additional term to represent the frequency-dependent transmission of the Faraday filter; since the measurements used elsewhere in this dissertation were collected at night, this term is ignored.

The total absorption cross section for the sodium atoms can be written in terms of the total absorbtion cross section (She, et al. 1992):

$$\sigma_{sa} = (g_2 / g_1)(\lambda_0^2 / 8\pi) A_{21} g(\nu). \qquad \text{(B.3)}$$

By combining equations (B.2) and (B.3), along with the Doppler broadened and Doppler shifted lineshape, the backscattered photons can be written in terms of the temperature and LOS wind.

However, equation (B.2) also depends upon other factors besides temperature and LOS wind. After the background photon counts are subtracted, power fluctuations and lower atmospheric fluctuations can be eliminated by normalizing the returns to Rayleigh scattering between 20 and 40 km. Rayleigh-Mie scattering is assumed to be independent of frequency over the narrow bandwidth of the sodium hyperfine transition. Therefore, the differences in the scattered returns between the three frequencies from the lower atmosphere can be attributed to differences in the laser power for each transmitted frequency, and differences in the normalized results from the MLT region for the three frequencies can be attributed to differences in the temperature/wind dependent backscattering cross-section and attenuation through the sodium layers.. Two ratios of the sodium layer returns can be formed, one of which is more sensitive to temperature and the other which is more sensitive to LOS wind. For the CSU Na lidar, these ratios are defined as:

$$R_T(T,V) = \frac{N_+ + N_-}{2N_0}, \tag{B.4}$$

which is more sensitive to changes in temperature, and

$$R_W(T,V) = \frac{N_+ - N_-}{N_0}, \tag{B.5}$$

which is more sensitive to changes in LOS wind. N_0, N_+, and N_- are background-subtracted, Rayleigh-normalized photon counts at each laser frequency. Figure B.2 shows a calibration curve which is used to convert these ratios into values of temperature and LOS wind. Assuming zero vertical wind, LOS winds can easily be converted to horizontal (meridional and zonal components) by correcting for the angle of the laser beam from zenith.

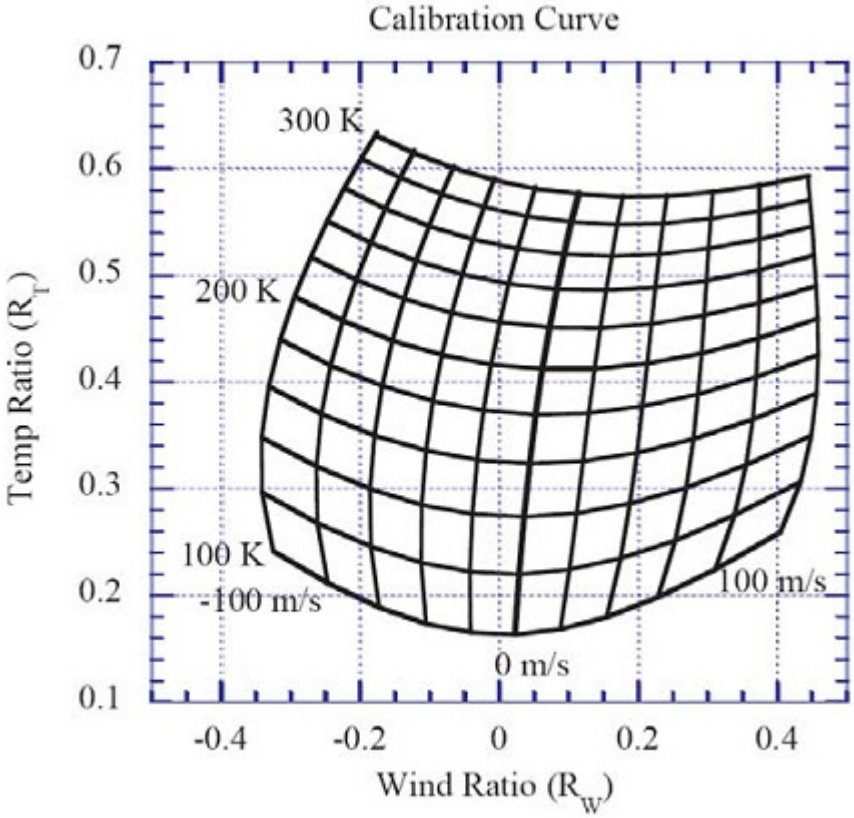

Figure B.2: Calibration curves for relating R_W and R_T to temperature and LOS velocities.

B.3 Acousto-optic modulation for frequency shifting

The laser is now locked to high precision at the D_{2a} peak. The two other frequencies needed to probe the sodium spectrum are produced using a dual-pass acousto-optic modulator (AOM). This device uses acousto-optic crystals, which use piezo-electric transducers set at 315 MHz. This causes the crystal to vibrate at this frequency. An incident beam of light will be shifted by the frequency of this traveling acoustic wave, and it will also be shifted in direction. To understand this interaction, the particle picture of both the light and acoustic wave can be invoked—the light is a photon with momentum proportional to the wavevector $\mathbf{k}_l$ and the acoustic wave is a phonon with momentum proportional to its wavevector $\mathbf{k}_a$. When an incident photon is struck by a phonon, the phonon will transfer its momentum and energy to the photon. The conservation of momentum, proportional to $\mathbf{k}_{l\pm}= \mathbf{k}_l \pm \mathbf{k}_a$, will lead to a change in direction of the exiting photon, while the conservation of energy, proportional to $\nu_{\pm}= \nu_0 \pm \nu_a$, will lead to a change in frequency of the photon.

The optical layout of the AOM is pictured in figure B.3. It is important for the light to exit the AOM with the same light path regardless of frequency; thus, there are optical elements chosen to correct the direction of the light. There is also a three-slot optical chopper rotating at 50 Hz. The holes are cut in different radial positions on the wheel—two are positioned to pass the shifted beams, while one passes the ν_0 (undeflected) beam. The chopper is necessary because the crystals have an efficiency of around 60%; the slotted chopper is needed to block the unwanted ν_0 beam from being transmitted while the crystal is on. Figure B.3 presents the three cases with light entering from the left, and propagating to the right. In case A, with both crystals off, the light

propagates straight through the two crystals, through the chopper, reflects off of a mirror, and then propagates back along the same path, providing light at ν_0. For case B, with the -315 MHz crystal (crystal 1) on, the light will propagate through that crystal and will be shifted in both frequency and direction. The presence of the focusing lens following the crystal redirects the light to pass through the second crystal as indicated in figure B.3. After passing through the chopper, the ν_- beam is reflected off of the mirror, is focused by the lens back through crystal 2, then the light is again focused to pass at an angle through crystal 1, where it is again shifted by -315 MHz, for a total shift of -630 MHz. The crystal also redirects the light to exit the AOM along the same path as the ν_0 light. As can be seen in figure B.3, case C, the procedure is similar for crystal 2, which produces a +630 MHz total shift (+315 MHz per pass).

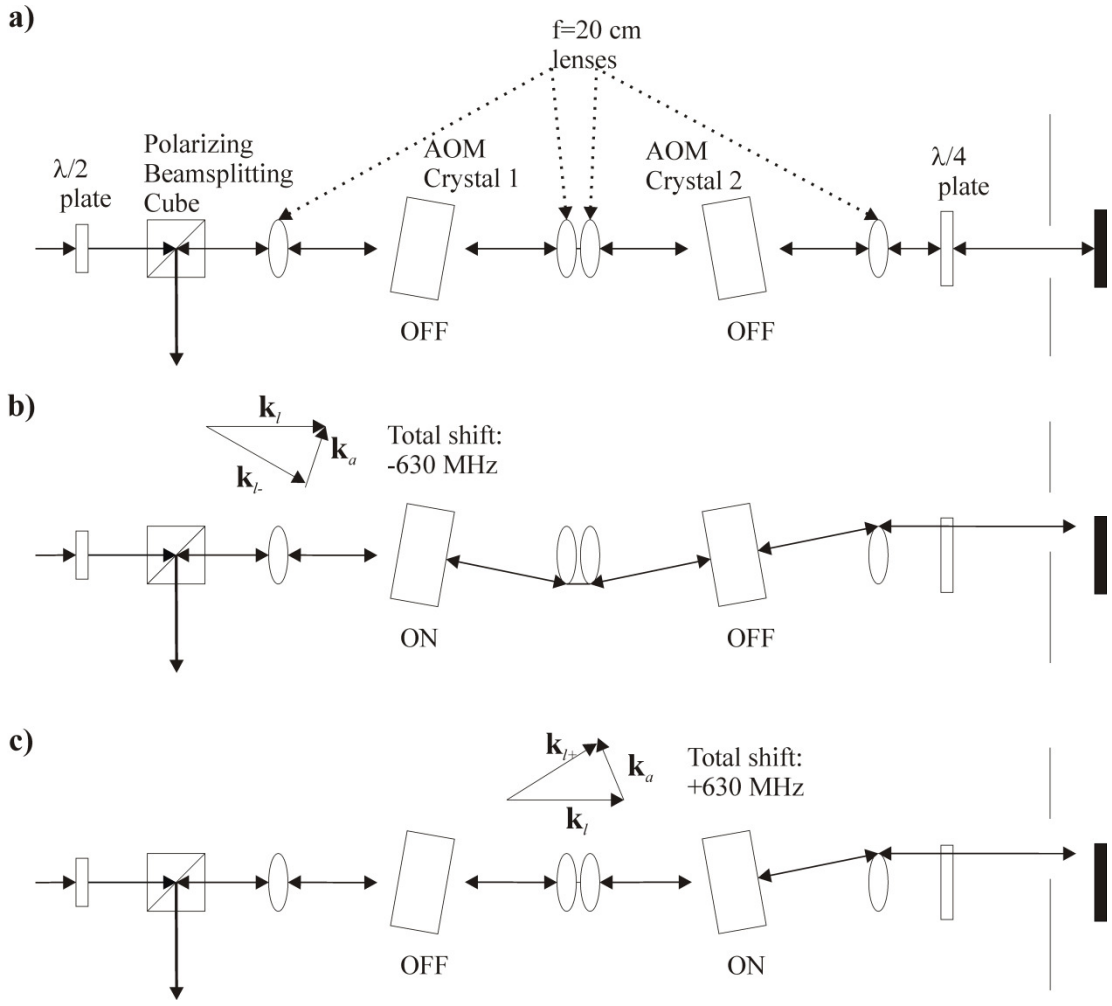

Figure B. 3: The dual-pass acousto-optic modulator optical layout. Part (a) shows the path of the unshifted beam passing at the D_{2a} peak. (b) and (c) show the ±630 MHz shifted beams.

Maintaining polarization is also important throughout the system. Due to a Brewster-angled dye jet and some optical elements in the ring dye laser, the output is linearly polarized. At the entrance to the AOM, the light passes straight through a polarizing beam splitter. Before the chopper, there is a quarter waveplate, which changes the linear polarization into a circular polarization. After reflecting off of the mirror, the light passes back through this waveplate and is linearly polarized, but orthogonal to the incident polarization. Thus, the light exiting the AOM reflects at the polarizing cube

beamsplitter, and has the proper polarization to pass through the remaining optical elements and be transmitted via a pulsed dye amplifier into the mesopause, and with an optical isolator between the c.w. and pulsed system, no reflected light can propagate back into the c.w. ring laser.

B.4 Pulse amplification and frequency chirp measurement

After the light has been processed with the optics of the AOM, it has a much lower power than the ring laser output, due to the many optical surfaces as well as the efficiency of the AOM crystals, requiring power amplification before transmission into the atmosphere. Also, in order to easily make time of flight measurements, it is necessary to transmit pulses of light. The CW beam is pulsed and amplified by the Pulsed Dye Amplifier (PDA).

The PDA (Spectra Physics QuantaRay PDA-1) uses high-energy pulses from a Nd:YAG laser (Spectra Physics QuantaRay PRO-230) to amplify the seed beam from the AOM in three successive dye cell amplifiers. This process takes the approximately 100 mW of CW seed light entering the PDA and amplifies it to an average pulse power of 1 W. This is enough power to transmit into the mesopause, scatter off of the sodium layer, and return with a high enough signal-to-noise ratio to be able to accurately determine temperatures to within ±1 K and winds within ±2 m/s.

The PDA amplification process, however, has a pulse linewidth of around 120 MHz, and it introduces slight, variable frequency shifts in the centroid of the pulse. This shift is mainly caused by the intensity dependence of the dye's refractive index, leading to a time-dependent gain of the of the PDA and in turn a time-dependent phase shift of

the amplified pulse.. Other external factors can contribute to the shifts, such as imperfect seeding of the Nd:YAG pump laser, and misalignment of the Nd:YAG pump and CW seed beams inside the PDA.

The frequency shift, which is referred to to as "chirp," will have an effect on the lidar determination of temperature and winds. However, the effect on temperature measurements is negligible, since a shift of 1 MHz causes an error of about 0.15 K, far smaller than the temperature error due to statistical photon noise. For wind measurements, the error induced is larger; a line of sight (LOS) wind bias of several meters per second is expected, since a frequency shift of 1 MHz is equivalent to a LOS wind bias of 0.6 m/s.

The unique solution developed at CSU is the wind bias monitor (Yuan, et al. 2009). This device utilizes a feature in the I_2 absorption spectrum which lies close to the D_{2a} peak. At the exit of the PDA, a beamsplitter reflects about 2% of the output into the iodine cell and an associated reference channel, as shown schematically in figure B.4. The light is detected by photodiodes. The iodine spectrum is shown in figure B.5, with the locations of the three PDA output pulses at ν_0, ν_+, and ν_-. A ratio can be formed, which will be sensitive to small shifts in the frequencies of these pulses:

$$R_c = \frac{T_- - T_0}{T_- + T_0} \tag{B.6}$$

where T_- is the background-subtracted, power normalized transmission through the iodine cell for ν_-, and T_0 is the same for ν_0. These are recorded experimentally, concurrent with lidar operation. To compare R_c values to $\Delta\nu_c$, the frequency shift of the output signal, the measured R_c is compared to a theoretically-generated calibration curve of R_c vs. $\Delta\nu_c$. This

curve is determined from the correlation of the pulsed lineshape function and the c.w. iodine transmission spectrum, which are both measured in the laboratory. For a range of reasonable $\Delta\nu_c$ values, we create a third-order polynomial fit of this calibration curve to use in the data analysis procedure to convert the measured values of the chirp ratio to frequency offset values, which correspond to LOS velocity changes as $\Delta V_c = -(589.158nm)\Delta\nu_c$. This system shown in Figures B.4. and B.5 has been shown in Yuan et al. (2009) to reduce the chirp-induced bias in the LOS wind measurements from around -5 m/s to less than -1 m/s.

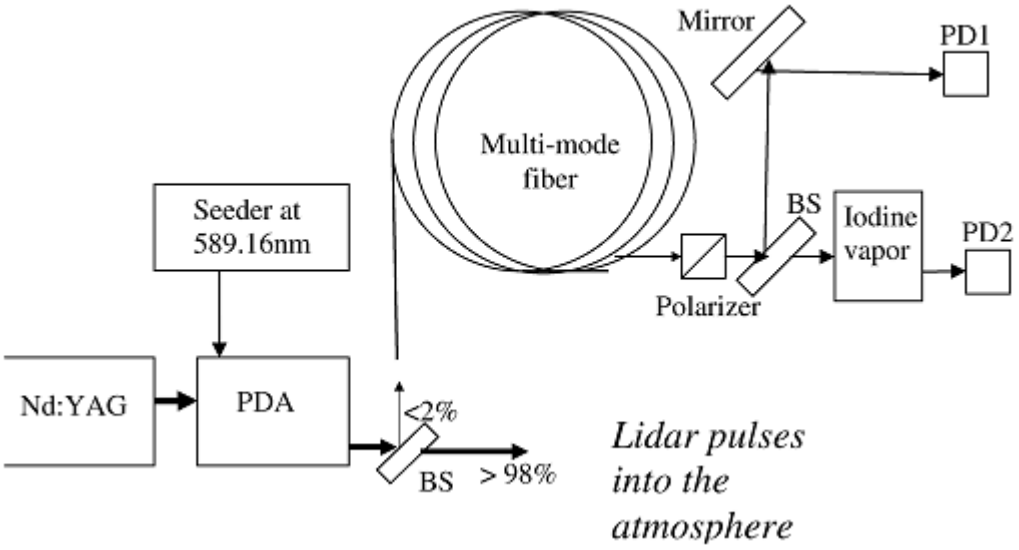

Figure B.4: The chirp subsystem. PD is photodiode, BS is beamsplitter. From Yuan et al. (2009).

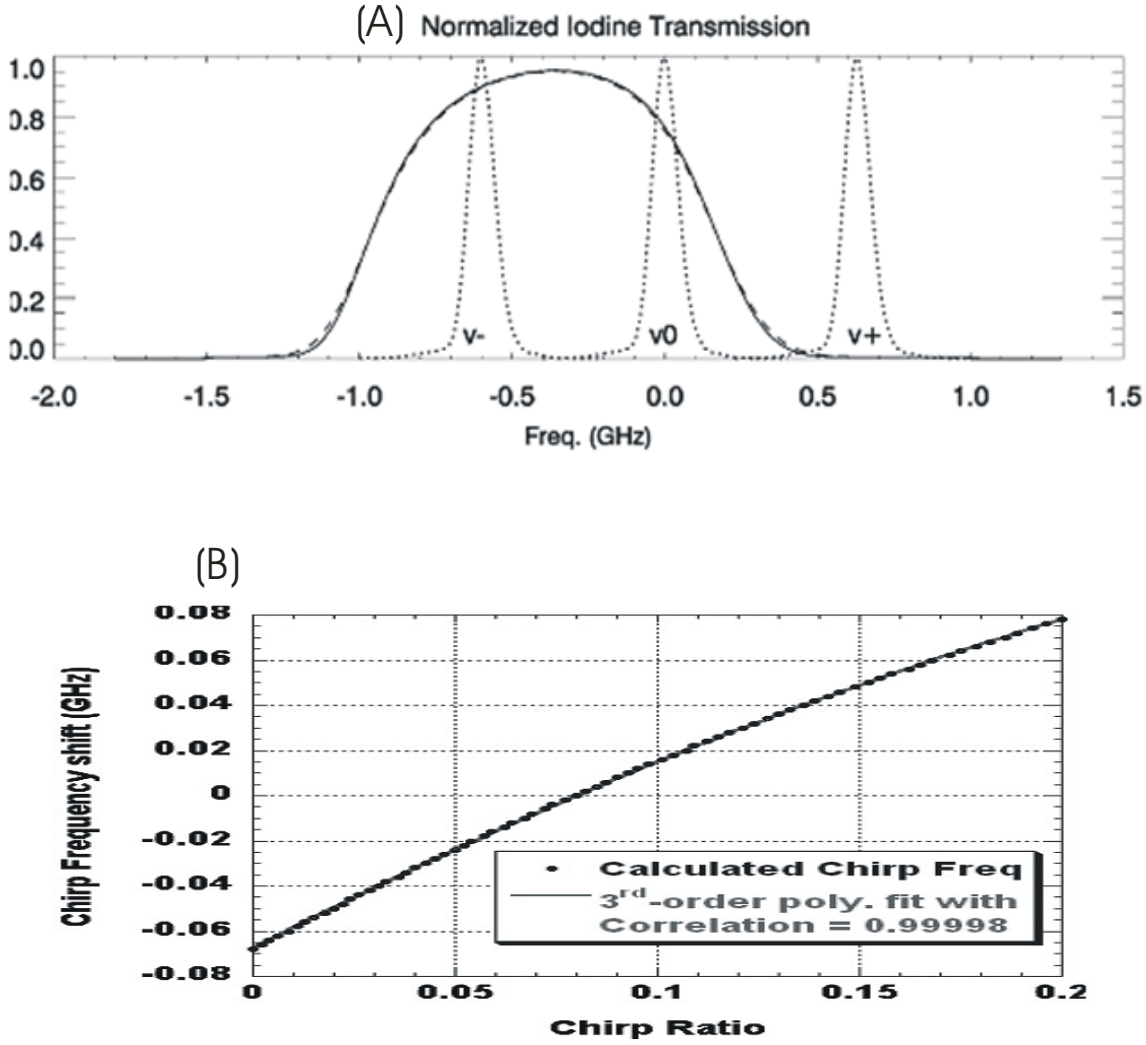

Figure B.5: The chirp transmission and calibration curve. (A) Transmission function: the solid line is the I_2 spectrum when probed with the CW laser, the dashed line is the spectrum when probed with the PDA output pulses, and the dotted line shows the relative placement and width of the three-frequency PDA output pulses. (B) The polynomial fit to the chirp frequency shift (GHz) vs. chirp ratio R_c.

B.5 The lidar receiver

The lidar receiver system, while in many ways conceptually simpler than the transmitter, still includes one important technical innovation (which is also central to the FFBS): the Faraday filter. Other sub-systems include the telescopes and fiber coupling,

which collect the light, and the PMT detectors which convert the returned photons to electrical signals for data recording.

B.5.1 The telescopes and fiber coupling

The light scattered from the Na atoms in the mesopause region is collected by telescopes located on the ground. Currently, there are three telescopes available to lidar detection: one Celestron 14" Schmidt-Cassegrain telescope, and two 30" Starsplitter Newtonian telescopes. The larger Newtonian telescopes allow for a factor of 5 improvement in the signal, and allow for the determination of nighttime winter zonal momentum flux.

Since the divergence of the lidar beam is small (1 mrad full angle), this means that the étendue of the telescope is small enough to allow the light to be fiber-coupled into the photomultiplier tubes. This has several practical advantages to conventional optics, the most important being that it allows for placement of the PMT's some distance away from the telescope, as well as limiting the number of optics that need to be adjusted to achieve proper focusing of the returned signal onto the PMT's.

B.5.2 The Faraday filter for lidar use

Since the Faraday filter principles have been well documented elsewhere in this dissertation (Chapter 4), I will only explain their role in the lidar system briefly. The light from the fiber optics is roughly collimated using a single lens (the 0.37 NA, 1.5 mm core diameter fiber uses a 25 mm focal length achromat). At night, a 1 nm FWHM interference filter centered at the D_{2a} transmission is all that is required to decrease the

background signal to an acceptable level. During the daytime, however, the presence of the solar background requires a filter with a much narrower bandwith. This is achieved with a Faraday filter. Figure B.6 shows a typical Faraday filter transmission function used for lidar data taking. The "flat top" of the center peak of the transmission is used because it roughly equalizes the transmission for the 10 lidar return frequency channels (which are due to the hyperfine structure of sodium—see chapter 4). During the daytime, the Faraday filter is inserted in the beam path leading to the PMT. This has allowed for 24-hour continuous measurement of temperature and winds since 2002, and allows for studies of atmospheric solar tides and their influence on the climatology of the mesopause region.

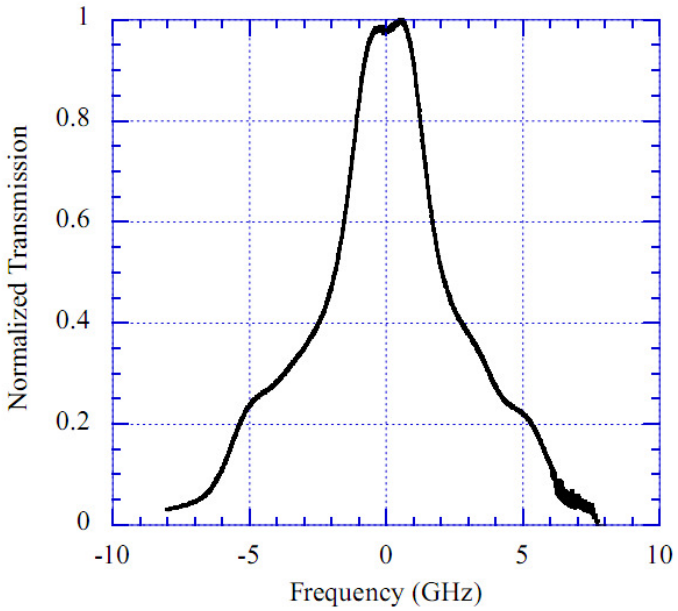

Figure B.6: Sample Faraday filter transmission for lidar. Notice that the transmission is scaled to unity at the peak for convinience; this is because the lidar data analysis only relies upon relative filter transmission.

B.5.3 Photon Detection

In either day or night, the lidar return signals are detected by PMT's. Currently, we have four total PMT's—two with 20% quantum efficiency (QE), and two with 40% QE—which are dedicated to lidar use. These PMT's are designed to count photons, i.e. to convert each photon detected into a measurable electronic pulse. To block strong lower atmosphere returns that would damage the PMT's, two mechanisms are employed. The 20% QE PMT's use a gating circuit, which switches the gain of the PMT to high voltage only after a certain time after the laser pulse was emitted. The 40% QE PMT's utilize a mechanical chopper wheel to block these lower atmosphere returns. These pulses are then collected by a counting card in a computer.

APPENDIX C

CURVE FITTING FOR DATA ANALYSIS

As presented in Chapter 7, the data as analyzed has large mean value variation and error bars for R_{D1}, R_{D2}, R_{DX}, and R_{DA}. This was shown to be due to low signal to noise ratios for the measured signals, and also possibly to geophysical variability within each 0.5 hour averaging period. Another possibility is to fit the data to a curve and use this fit to calculate R_{D1}, R_{D2}, R_{DX}, and R_{DA}. The curve will effectively smooth over the noise and other variability, hopefully yielding results more indicative of the actual values of these important nightglow parameters.

This appendix details the curve fit method used for this analysis. It also contains preliminary results of this analysis technique for one night of measurements and a discussion of the usefulness of this technique and suggestions for future improvement of the data analysis procedure.

C.1 Curve fit procedure

The raw data plotted in Chapter 7 (i.e., Fig 7.2) can essentially be broken up into straight line segments. For that reason, I chose to fit the nightglow data to the curve:

$$s_{fit} = A + Bt \qquad \text{C.1}$$

where A and B are the fit parameters (intercept and slope, respectively), t is the time, and s_{fit} is the fit value for one of the measured nightglow signals. The fit is performed by

minimizing the value of χ^2. With the parameters A and B determined, s_{fit} can be determined for the midpoint of each half hour time bin. After dark count subtraction, this is the value used for S as defined in eq. (5.2). The data processing can then proceed as described in section 5.3.1.

The error in s_{fit} is propagated from the error in A and B (ΔA and ΔB), which is determined from the fitting process, and includes the contribution from the measured signal error. From this, the error in S and the nightglow intensities I and the ratios can be calculated using equations 5.10-5.15.

The procedure described in this section could also be performed using other fits, for example, a fourth-order polynomial, which would capture more of the short-term variability in the data.

C.2 Results from 26 September 2009

One night of data analyzed in the manner described in section C.1 is presented here. The data from 26 September 2009, shown previously in Figure 7.2, can be fit with two lines: one for the data before 6 UT, and another for the data after 6 UT. The results are shown in figure C.1. Table C.1 lists the fitting parameters and errors for the two fits.

Several features can be seen by looking at the results in Figure C.1. Where I values are very small and above and below zero, the absolute average values of R_{D1}, R_{D2}, R_{DX}, and R_{DA} become quite large and have larger error bars. This is true during the first half of the night. During the second half of the night, the values are significantly more constant since all of the I values are positive. For R_D data after 6 UT, the value is nearly constant; the reason for this is not clear.

Table C.1: Fitting parameters and errors for the data from 26 September 2009.

fit		A	B	ΔA	ΔB	χ^2
2 to 6 UT	S_{A1}	10274.3	-738.8	35.2	8.1	2525.3
	S_{B1}	5168.0	-390.9	24.6	5.6	1325.1
	S_{A2}	6961.48	-416.6	30.2	7.0	4196.9
	S_{B2}	12275.6	-1210.0	35.6	8.04	3970.4
6 to 10 UT	S_{A1}	4561.8	265.4	57.7	7.2	671.5
	S_{B1}	1899.1	184.9	40.9	5.1	248.8
	S_{A2}	4256.3	76.8	49.6	6.2	500.1
	S_{B2}	2243.8	583.3	57.8	7.3	501.7

C.3 Discussion

Overall, the error bars in Fig. C.1 are larger than the data analysis presented in Chapter 7 since the χ^2 and error values in the fit parameters in Table C.1 are fairly large. This indicates that the data has variation besides that due to photon noise—it could be possible to determine the contribution of this non-random error (Roberts 2010). The χ^2 values also indicate that the data for the second half of the night is more linear since the χ^2 values are smaller than those for the first half of the night.

Also as shown in Fig. C.1, some values of I are small and of opposite signs and this causes the mean values of R_{D1}, R_{D2}, R_{DX}, and R_{DA} to increase rapidly. For intervals where this is the case, neither the curve fit method described here nor the original data analysis presented in Chapters 5 and 7 appear to work. This all suggests that the best way

to improve the R_{D1}, R_{D2}, R_{DX}, and R_{DA} data sets are to optimize the temperature settings of the Faraday filters for minimum error in the measurement of the four I values.

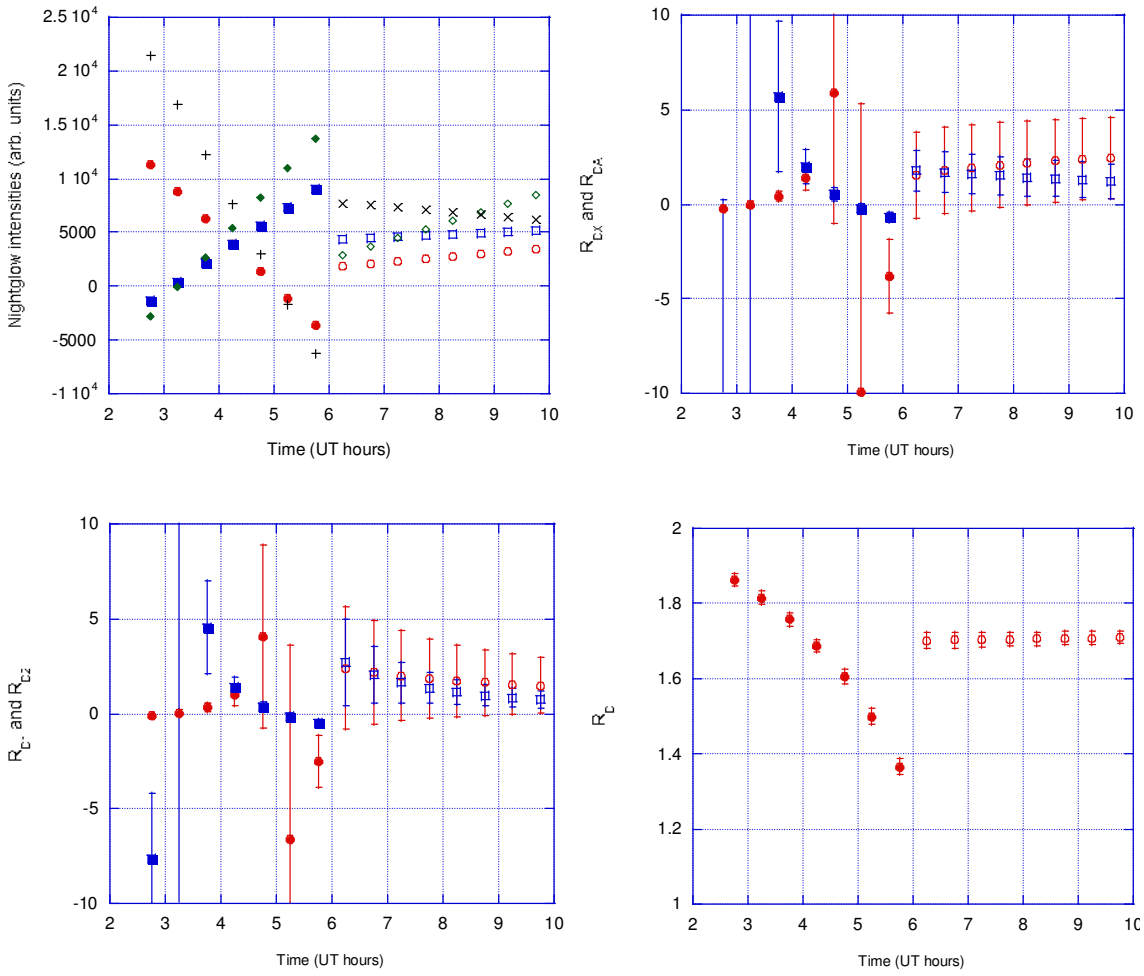

Figure C.1: 26 September 2009 data analyzed with the linear curve fit method. A. Calculated nightglow intensities: Red circles: I_{1X}, Blue squares: $I_{1A,}$ Green diamonds: I_{2X}, Black crosses: I_{2A}. B. R_{DX} (Red circles) and R_{DA} (Blue squares). C. R_{D1} (Red circles) and R_{D2} (Blue squares). D. R_D. Data from 0.5 hour intervals. Solid points are the fit before 6 UT, and open points are after 6 UT. Error bars are calculated by propagating the error in the fitting parameters.

In the future, this curve fit method could provide a better estimation of the nightglow ratios than the method presented in Chapter 5. This is due to the smoothing effect of the fitted curve. For this method to work, it would be necessary to determine the fraction of the variability in the measured signals due to photon noise, as opposed to the

fraction due to non-random variations. Also, other curve fits besides the linear fit might be better suited for these nightglow data sets—for example, higher-order polynomial fits may better account for some of the variation over the night than the linear fit. Although the fitting method has the potential to improve the accuracy of the analysis, the results shown in Fig. C.1 do not show a clear improvement. The improvement would be more reliable and robust if the FFBS is redesigned to increase the Faraday filter transmission and thus the signal to noise ratio of the low signal channel.