NONPARAMETRIC AND PARAMETRIC METHODS FOR
SOLAR OSCillation SPECTRA

by

CHARLOTTE A. L. HALEY

A thesis submitted to the
Department of Mathematics and Statistics
in conformity with the requirements for
the degree of Doctor of Philosophy

Queen’s University
Kingston, Ontario, Canada
September 2014

Copyright © Charlotte A. L. Haley, 2014
Abstract

The study of the systematic oscillations of the Sun has led to better understanding of the Sun’s inner structure and dynamics, and may help to resolve inconsistencies between observations and the standard solar model. Recent studies have concluded that solar modal structure remains coherent past turbulence in the convection zone and imprints its signatures on the solar wind and the interplanetary magnetic field fluctuations, and these structures are coherent with atmospheric pressure variations, terrestrial seismic oscillations, and data from communications systems. Time series containing modal structure can be expected to contain several thousands of resolved and unresolved line components in very short bands in frequency, and the measurement of these modes pushes spectrum estimation methods for time series to its limit.

This thesis presents two theoretical contributions for modeling solar oscillations in power spectra (i) expressions for the expected number and shape of significant spurious peaks in spectrum estimates are given, in the absence of modal structure, and a permutation test for the identification of spectra containing pathological numbers of modal components. (ii) A model for maximum likelihood estimation of the solar oscillation parameters in composite spectra is given. The scientific contributions of this thesis are (a) identification of highly significant modal artifacts in solar wind measurements as seen by the Advanced Composition Explorer (ACE) on the 2 – 3mHz band and (b) quantification of the presence of modal structure in secondary cosmic rays (specifically neutrons) on Earth.
Co-Authorship

The contents of Chapter 3 are co-authored with David J. Thomson.
Acknowledgements

My sincere thanks to David Thomson for countless hours of academic guidance, excellent conversations, and teaching me the courage of my convictions (mainly by suggesting controversial topics for me to defend at meetings). And to Maja-Lisa Thomson for mentoring me while making a mentor of me at math camp.

I’d also like to thank professors Selim Akl and Damian Redfearn for all of their academic help and encouragement during my time in graduate school.

To my colleagues from Queen’s, Carly Rozins, Johanna Hansen, Asia Matthews, Wes Maciejewski, Gina Hochban, Caroline Seguin, Karim Rahim, Wesley Burr, Aaron Springford, Gwen Eadie, Josh Pohlkamp-Hartt, Dave Riegert, Sami Torbey and Geoff Seaborn, I wish you the very best. Thank you for making my time at Queen’s memorable.

But mainly this thesis is for my family, Susan Haley, Anna Haley, and Marten Hartwell, for their love and support. And for Dustin West, for being a fountain of hilarity and adventure throughout.
In memory of Marten Hartwell.
Contents

Abstract i

Co-Authorship ii

Acknowledgements iii

Contents v

List of Tables ix

List of Figures x

Chapter 1: Introduction 1

1.1 Organization ................................................. 2

1.2 Note ........................................................... 3

1.3 Data Sources and Acknowledgments ......................... 4

1.4 What’s new? .................................................. 5

Chapter 2: Solar Physics and Helioseismology 6

2.1 The Sun ....................................................... 7

2.2 Elementary Helioseismology ................................ 12

2.3 Why Helioseismology? ........................................ 15

2.4 Literature Review ............................................ 16

2.4.1 The Discovery of the Solar Modes ..................... 16

2.5 Modeling of Solar Oscillations ............................. 17

2.5.1 Further Observations .................................... 19
Chapter 3: Data

3.1 Optical Doppler measurements .................................................. 22
3.2 Solar Wind Data ........................................................................... 24
3.3 Cosmic Ray Data .......................................................................... 30
   3.3.1 Neutron Data ........................................................................ 31
3.4 Conclusions .................................................................................. 40

Chapter 4: Spacing and Shape of Peaks .................................................. 43

4.1 Crossing Problems ........................................................................ 45
   4.1.1 The Rice Equation ................................................................ 46
   4.1.2 Extension to $\chi^2$ Processes ................................................. 47
   4.1.3 Direct Spectrum Upcrossing Rate ............................................ 49
   4.1.4 Multitaper Antecorrelation ....................................................... 50
   4.1.5 Multitaper Spectrum Upcrossing Rate ...................................... 53
   4.1.6 Simulation .............................................................................. 54
4.2 Shape and Width of Spurious Peaks ............................................... 56
   4.2.1 Width of Spurious Peaks ......................................................... 56
   4.2.2 Shape of Spurious Peaks ......................................................... 57
4.3 Example ......................................................................................... 60
   4.3.1 ACE Proton Density ............................................................... 60
   4.3.2 Example: Neutron Data ......................................................... 62
   4.3.3 Example: Surgical Scheduling ................................................. 63
4.4 Extensions ...................................................................................... 64
   4.4.1 Harmonic F-test ................................................................... 64
4.5 Solar Modes and Multiple Comparisons .......................................... 66
4.6 Conclusion .................................................................................... 68

Chapter 5: Parametric Methods for Many Modes .................................... 70

5.1 Review: Modeling Solar Oscillations ............................................. 71
   5.1.1 Historical Notes .................................................................... 71
   5.1.2 M-averaged spectra ............................................................... 74
   5.1.3 Spectral Line Asymmetry ....................................................... 75
   5.1.4 Summary .............................................................................. 78
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.2 Signal Estimation and Detection</td>
<td>79</td>
</tr>
<tr>
<td>5.3 Eigenvalue Representations</td>
<td>81</td>
</tr>
<tr>
<td>5.4 Parameter Estimation</td>
<td>86</td>
</tr>
<tr>
<td>5.4.1 Partial Derivatives</td>
<td>86</td>
</tr>
<tr>
<td>5.4.2 Advanced Optimization</td>
<td>89</td>
</tr>
<tr>
<td>5.5 Practical Considerations for Real Data</td>
<td>89</td>
</tr>
<tr>
<td>5.6 Examples</td>
<td>90</td>
</tr>
<tr>
<td>5.6.1 Lorentzian Example</td>
<td>90</td>
</tr>
<tr>
<td>5.6.2 Seismic Data</td>
<td>94</td>
</tr>
<tr>
<td>5.7 Summary</td>
<td>95</td>
</tr>
<tr>
<td>Chapter 6: Summary and Conclusions</td>
<td>97</td>
</tr>
<tr>
<td>6.1 Summary</td>
<td>97</td>
</tr>
<tr>
<td>6.2 Future Work</td>
<td>99</td>
</tr>
<tr>
<td>6.2.1 Part I: Nonparametric Methods</td>
<td>99</td>
</tr>
<tr>
<td>6.2.2 Part II: Parametric Methods</td>
<td>100</td>
</tr>
<tr>
<td>6.2.3 Future Data Analysis</td>
<td>104</td>
</tr>
<tr>
<td>6.3 Conclusion</td>
<td>107</td>
</tr>
<tr>
<td>Appendix A: Nonparametric Spectrum Estimation</td>
<td>124</td>
</tr>
<tr>
<td>A.1 Direct Spectrum Estimators</td>
<td>124</td>
</tr>
<tr>
<td>A.2 Multitaper Spectrum Estimators</td>
<td>126</td>
</tr>
<tr>
<td>A.3 Standardized Spectrum</td>
<td>129</td>
</tr>
<tr>
<td>A.4 Rayleigh Resolution</td>
<td>130</td>
</tr>
<tr>
<td>A.5 Additional Tools</td>
<td>130</td>
</tr>
<tr>
<td>A.5.1 Harmonic F-test</td>
<td>131</td>
</tr>
<tr>
<td>A.5.2 Magnitude Squared Coherence</td>
<td>132</td>
</tr>
<tr>
<td>A.5.3 Complex Demodulation</td>
<td>133</td>
</tr>
<tr>
<td>Appendix B: Standard Tables of Upcrossing Rates</td>
<td>135</td>
</tr>
<tr>
<td>Appendix C: Karhunen Loève Expansion</td>
<td>139</td>
</tr>
<tr>
<td>Appendix D: Theorems Related to Eigenvalue Problems</td>
<td>141</td>
</tr>
</tbody>
</table>
List of Tables

<table>
<thead>
<tr>
<th>Table</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Summary statistics for the neutron data</td>
<td>35</td>
</tr>
<tr>
<td>3.2</td>
<td>Distribution of the neutron monitor spectra on 2-3mHz</td>
<td>39</td>
</tr>
<tr>
<td>4.1</td>
<td>Number of significant upcrossings observed and expected in ACE proton density spectrum</td>
<td>61</td>
</tr>
<tr>
<td>4.2</td>
<td>Neutron data permutation test 2-3mHz</td>
<td>62</td>
</tr>
<tr>
<td>B.1</td>
<td>Computed values for the constant $\psi$, Eqn 4.19</td>
<td>136</td>
</tr>
<tr>
<td>B.2</td>
<td>Expected number of upcrossings and dwell bandwidths for common choices of $NW$</td>
<td>137</td>
</tr>
<tr>
<td>B.3</td>
<td>Expected number of upcrossings and dwell bandwidths for common choices of $NW$</td>
<td>138</td>
</tr>
</tbody>
</table>
List of Figures

2.1 Cosmic ray variations and the sunspot cycle .......................... 10
2.2 Node circles of spherical harmonics ................................. 11

3.1 Optical Doppler spectrum and \((l, \nu)\) diagram ......................... 23
3.2 ACE solar wind data ........................................ 25
3.3 ACE standardized spectrum ...................................... 26
3.4 Decomposition of multitaper spectrum - simulation .................. 27
3.5 Mixture model for the distribution of the spectrum of solar wind data 29
3.6 Neutron monitor data from Thule, Newark, and South Pole .......... 33
3.7 Neutron monitor spectra from Thule, Newark, and South Pole ...... 34
3.8 Neutron monitor spectra contain large numbers of peaks .......... 36
3.9 Decomposition of neutron monitor spectra .......................... 38
3.10 Examples of splittings near mode frequencies in neutron data .... 40

4.1 Upcrossing rate and dwell bandwidth for MT and periodogram estimates 55
4.2 Average shape of spurious peaks in MT estimates of white noise ... 58
4.3 Average shape of spurious detections in periodograms of white noise 59
4.4 Simulation - number of spurious peaks in the F-test and MSC .... 66

5.1 Lorentzian transfer function profiles ............................... 72
5.2 Eigenvalues and Eigenfunctions of the Oscillating System .......... 85
5.3 Eigenvalues and eigencoeficients for simulated Lorentzian process .. 91
5.4 Eigenvalues and eigencoeficients of CDM of Lorentzian process .... 92

6.1 MSC crossing rates - simulation ................................. 100
6.2 LVF evolutive spectrum for muon data ............................. 105
Chapter 1

Introduction

This thesis is an attempt to introduce some useful nonparametric and parametric statistical methods for the purposes of signal detection of a particular physical process, namely the solar oscillations or “modes” which manifest themselves in the magnetic variations observed in interplanetary space and on Earth. The scientific findings of this thesis support the hypothesis that solar modal structure modulates the galactic cosmic radiation reaching the Earth from space, while the novel statistical methods can be broadly categorized into nonparametric and parametric methods.

In the nonparametric category, a description of the statistical properties of spectrum estimates containing *many lines* is sought, mainly the distributions of such spectrum estimators and their derivatives in frequency. Viewing the prewhitened spectrum estimator as an approximately stationary process, I will derive the number of significant upcrossings expected in spectrum estimates of Gaussian white noise, that is, those which contain no modal structure, as well as the width and shape of such significant excursions. This gives the rate of false detection of line components in spectrum estimators when none are actually present - a valuable piece of information when considering processes containing modal structure. By means of a permutation
test, I will show that the observed modal structure, that is, the large number of perceived line components in spectrum estimates of cosmic ray data is a direct consequence of the correlation structure of the process, and not the distribution of the data. Temporally permuted versions of the data possess spectrum estimates having the number of significant peaks expected were there no solar modes, which is in stark contrast with the large numbers of peaks actually observed.

In the parametric category in this thesis, one begins with a spectrum estimate on data obtained from images of the solar disk and attempts to estimate the center frequency, quality factor, and amplitude of each mode of interest, if the mode is present. This is referred to as peak bagging in the literature, a term coined by Jesper Shou, an avid climber and solar physicist. I will describe the mainstream methods in use for this estimation problem, and develop a maximum-likelihood technique to determine modal parameters. Because the covariance structure particular to harmonic oscillators has eigenfunctions and eigenvalues which can be determined analytically, the likelihood has a particularly convenient form.

The two methods halves of this thesis are joined with numerous examples of space physics processes which appear to contain modes, namely, heliosesmic data (in which modes are mainly observed), solar wind data from ACE, and neutron monitor data from Earth.

1.1 Organization

This thesis contains two background chapters. Chapter 2 contains background information on solar physics and helioseismology, the physical characteristics of the Sun, a brief description of the standard solar model, and an introduction to the remote
study of the Sun’s seismic structure. Some knowledge of spectrum estimation is assumed for the data analysis in Chapter 3 and these details are given in Appendix A. The appendix also describes some nonstandard properties of the multitaper spectrum which are essential to the discussion in Chapters 3 and 4. In Chapter 3 I will motivate the problems at hand with real data, namely (i) optical Doppler measurements collected for the purpose of estimating helioseismic parameters (ii) solar wind data from interplanetary space (iii) neutron measurements from ground based instruments collecting data on charged particle influx from galactic sources.

The following two chapters give the theoretical results of the thesis along with examples. Chapter 4 is devoted to the spacing and shape of peaks in nonparametric spectrum estimates in the absence of modal structure [117]. I will derive expressions for the number of significant peaks exceeding a given significance level in a spectrum and verify these using simulations. Because estimation of these quantities can be computationally intense, Appendix B gives tables of crossing rates for various choices of the multitaper parameters (a concise version of the supplement to [117]). Examples using real data are also given. Chapter 5 is devoted to maximum likelihood estimation of the parameters of the harmonic oscillators that drive the solar oscillations. This model is fit to simulated data.

1.2 Note

Whether it is to the detriment of this thesis or not, the careful analysis of helioseismic time series is fundamentally a problem in physics. As such, this thesis, which is really about statistical signal detection in data with particularly harsh statistical properties, where spectra can be expected to contain several thousands or millions of modes in
relatively short bands in frequency, may seem to the statistician as excessively data-oriented or ad-hoc. There is a long preliminary chapter in which (albeit original) data analysis is presented for the purpose of motivating the statistical advances. However, the emphasis on the data is not here misplaced. It may be difficult to identify with a situation seemingly so far-fetched as the identification of perverse numbers of modes in time series. So many modes, in fact, that the identification is often done by quantifying the overall distribution of the spectrum, as opposed to focusing on individual line components. But without this motivation, the theoretical contributions of this thesis appear to have no application whatsoever. Reviewers have commented that the approach to quantifying false detection rates in signal detection problems in Chapter 4 is an unconventional solution to an unconventional problem, and this is the motivation for presenting the problem data–first.

1.3 Data Sources and Acknowledgments

International monthly mean sunspot numbers were obtained from the NOAA website http://www.ngdc.noaa.gov/stp/space-weather/solar-data/solar-indices/sunspot-numbers/international/tables/table_international-sunspot-numbers_monthly.txt.

The Global Oscillation Network Group (GONG), which produces near real time solar disk images and Doppler measurements, provided the 36-day global helioseismic time series. These data are readily available from http://gong2.nso.edu/archive/patch.pl?menutype=t

Ten-second neutron monitor data for this study was kindly provided by Roger Pyle from the Bartol Research Institute. The hourly values were obtained directly
1.4. WHAT’S NEW?

from the Bartol Research Institute ftp site, where the link is available from [http://neutronm.bartol.udel.edu/](http://neutronm.bartol.udel.edu/).

Hourly muon data from the Nagoya multidirectional muon telescope was obtained from [http://www.stelab.nagoya-u.ac.jp/ste-www1/div3/muon/muon1.html](http://www.stelab.nagoya-u.ac.jp/ste-www1/div3/muon/muon1.html).

Data sampled at 64-seconds from the SWEPAM instrument aboard ACE was obtained from [http://www.srl.caltech.edu/ACE/ASC/level2/lvl2DATA_SWEPAM.html](http://www.srl.caltech.edu/ACE/ASC/level2/lvl2DATA_SWEPAM.html).

1.4 What’s new?

Chapter 2 contains background material on helioseismology; none of these findings are new. Chapter 3 contains original data analyses, and the scientific conclusions I have come to in this chapter are new. Chapter 4 of this thesis is entirely new and has been co-authored with my thesis advisor, David Thomson, see [117]. These results have been discussed at the Joint Statistical Meetings [48]. Chapter 5 aside from the literature review, §5.1, is a new application of an old communications problem.
Chapter 2

Solar Physics and Helioseismology

Helioseismology is the study of global and local seismic oscillations of the solar interior. Inference based on oscillation frequencies determined using helioseismic methods has led to better knowledge of the structure of the solar interior, stellar evolution and composition, and better knowledge of properties of solar neutrinos which has the potential to contribute to our understanding of dark matter and particle physics. The success of helioseismology is built on a century of theoretical study of normal mode seismology on Earth.

The next section summarizes basic properties of the Sun. Section 2.2 is an introduction to solar oscillations, with a short literature review on their detection and observation. Section 2.3 summarizes some of the reasons helioseismic study is important, and how it has contributed in other fields. Some excellent reviews on helioseismology are those of Christensen-Dalsgaard [23], Narayanan [77], and the book of Stix [105]. The books by Aerts et al. [2] and Chaplin [22] are specially written for a lay audience.
2.1 The Sun

The Sun is a sphere of radius $r_\odot \approx 6.96 \times 10^8$ km (the subscript $\odot$ can be read as “of the Sun”), with mass $m_\odot = 1.989 \times 10^{30}$ kg. It is at an approximate distance of 1 astronomical unit (A.U. $\approx 1.50 \times 10^8$ km) from the Earth, depending on the time of year and is in a state of *hydrostatic equilibrium*, in which inwardly directed gravitational forces in the Sun are balanced by an inwards increase in pressure \[105\].

The Sun is currently within its *main sequence* state, in which hydrogen is converted to helium in the core, during which its luminosity ($L_\odot \approx 3.84 \times 10^{26}$ W, total energy output (photons), per unit time) and radius have increased ($\dot{r}_\odot \approx 2.4$ cm/year) \[104\], \[105\] Table P.2]. The Sun’s luminosity is related to the solar flux as luminosity per unit area (W/m$^2$).

Heat from nuclear fusion of hydrogen to helium in the core drives constant gravitational expansion and contraction \[51\]. This net reaction can be written as \[23\] Eqn. (4)

$$4^1\text{H} \rightarrow ^4\text{He} + 2e^+ + 2\nu_e,$$

(2.1)

where neutrinos, $\nu_e$, and positrons, $e^+$, are produced. It is this gradual change in composition which mainly drives solar evolution.

Solar Magnetic Activity

The Sun’s magnetic field has an 11-year period and reverses direction every cycle around the solar maximum, resulting in a 22-year magnetic, or *Hale cycle*. The Sun has been observed to follow the Hale cycle since 1909 \[47\]. Solar magnetic activity
comprises the sunspots, solar flares and solar wind variations. Sunspots are lower-temperature “dark” areas on the surface of the Sun, possessing strong magnetic fields (up to 0.4 Tesla, [23]) of their own which inhibit energy transport by convection from the interior to the surface. The sunspot numbers have been recorded since the 1600’s and their behavior captures the gross features of the 11-year solar cycle. Faculae, in contrast, are bright, hot, spots on the solar surface sometimes accompanying sunspots; when these occur in the chromosphere (middle surface layer of the Sun) they are called plages. Solar flares are large, bright releases of energy of up to $6 \times 10^{25}$ J, powered by a sudden release of magnetic energy issuing from the corona, the layer of hot plasma above the surface of the Sun.

Solar wind particles stream outward from the Sun with supersonic velocity (1.5 – 10 keV) into the interplanetary medium until they reach the termination shock, where the solar wind meets the interstellar medium. Due to differential rotation, i.e. faster rotation at the equator than at the poles, the Sun’s magnetic field lines are twisted, causing magnetic field loops and producing the solar dynamo. Understanding of the solar internal rotation sheds light (through helioseismology) on the mechanism which produces the dynamo.

Solar magnetic phenomena such as coronal mass ejections (CMEs), flares, and ground level enhancements (GLEs), in which charged particles reach Earth at high speeds, perturb magnetic fields in ways which can harm orbiting satellites and interfere with power grids and communications systems.

The heliosphere is the outer atmosphere of the Sun and marks the edge of the Sun’s magnetic influence in space. The solar wind shields our solar system from the majority of galactic cosmic radiation (high energy charged particles originating in
2.1. THE SUN

Cosmic ray intensity is seen to vary with the 11-year solar cycle and the 22-year magnetic cycle, see Fig 2.1. The 11-year relationship is striking upon visual inspection, though there appears to be a brief time lag in the neutron data. Fig. 2.1 shows monthly average sunspot numbers (see the NOAA website) compared with Thule neutron multiplicity (counts per hour averaged every 27 days) over the course of a few solar cycles. The sunspot series is here a convenient proxy for the level of solar magnetic activity. Fig. 2.1 also shows muon data from the Nagoya muon telescope (bottom plot), for comparison. Neutron monitor data and muon monitor data are proxies for the level of secondary cosmic radiation, about which more will be said in the next chapter.

The periodic variation with the sunspot cycle aside, non-periodic variation due to solar transient events is also a significant contributor to variations in cosmic ray intensity. These solar effects include (i) Forbush decreases, which are sharp drops in cosmic ray intensity due to coronal mass ejections (CMEs), or violent bursts of magnetic charge issuing from the corona. CMEs can be accompanied by (ii) ground level enhancements (GLEs), which occur when solar radiation reaches energies registered by neutron monitors. (iii) Solar flares have the effect of increasing intensity because of an influx of low energy particles and are sometimes accompanied by CMEs.

Standard Solar Model

The standard solar model is a set of differential equations and boundary conditions describing the Sun’s structure as if it were a spherically symmetric main sequence star.

\[ \text{http://www.stelab.nagoya-u.ac.jp/ste-www1/div3/muon/} \] The muon telescope has been operational since 1970.
2.1. THE SUN

The differential equations describe the pressure, mass, temperature, and luminosity of the Sun as a function of radius with respect to time, see [23] Eqns (2). It is out of the scope of this review to discuss the subtleties of this representation, but this model describes two means of energy transport in the Sun, namely convection and radiation, [105].

Figure 2.1: Prominent 11-year cycles are visible in the sunspot series (a), (units are monthly number of sunspots), and the neutron monitor data from Thule (b), (reported in counts per hour averaged over 27-day periods), where peaks in the solar data correspond to troughs in the cosmic ray incidence on Earth. This effect is due to the deflection of cosmic radiation away from the Earth when solar magnetic activity is high [80]. Muon relative intensity (c) from the vertical component of the multidirectional muon telescope at Nagoya, Japan are also shown. Though there is a strong yearly periodic trend, and long-term upwards trend, muon variations appear consistent with the neutrons.
2.1. THE SUN

Gross Features of the Sun’s Internal Structure

Figure 2.2: Panel (a) shows gross features of the Sun’s internal structure. The top spheres in panel (b) show node circles of spherical harmonics and are labeled for the various choices of \((l,m)\) they represent, while the shapes directly below show the global oscillations as if they were in motion. Note the leftmost mode is a fundamental mode. Mode amplitude, which here corresponds to velocity measurements on the surface, has been greatly exaggerated for illustration purposes. Surface velocity measurements per \(p\)-mode can achieve a maximum velocity of about 20cm/s and can resonate for a few hours (for high \(l\) \(p\)-modes) to several months (low \(l\) \(p\)-modes), [51].

Internal Structure

Traveling outwards from the Sun’s center, the core is where the bulk of the Sun’s energy is generated, which has radius \(r_\odot/4\). Moving outward from the core, there is a radiative zone where radiative diffusion is the mechanism for energy transportation. The tachocline represents the boundary (at 0.713\(r_\odot\)) between the radiative and the convective zone. In the convective zone, thermal convection is the primary mechanism for energy transport, and interactions there are considered adiabatic (heat does not enter or leave the system). The left panel of Fig. 2.2 summarizes these features. The Sun’s atmosphere consists of the photosphere, the optical surface of the sun, the middle chromosphere layer, and the corona.
2.2 Elementary Helioseismology

The goal of helioseismology is to observe oscillations on the solar surface so as to infer information on the solar interior such as the internal rotation rate. Perturbations of the Sun away from hydrostatic equilibrium cause the Sun to oscillate, much like the Earth following an earthquake, and there are two restoring forces for these oscillations, pressure and gravity. For an excellent reference on asteroseismology, consult [2].

Beginning with basic hydrodynamic differential equations and boundary conditions described in [23] §V.A, the application of perturbation theory yields an eigenvalue problem, whose solutions are oscillations coming from a discrete set of frequencies. These (separable) solutions are the spherical harmonics, $Y_{lm}^m(\theta, \phi)$, defined as

$$Y_{lm}^m(\theta, \phi) = (-1)^m c_{l,m} P_l^m(\cos \theta) \exp (im\psi)$$

(2.2)

where $P_l^m$ are the Legendre functions, $c_{l,m}$ is a normalization constant so that the $L^2$-norm of $Y_{lm}^m(\theta, \phi)$ is unity, $\phi$ is solar longitude and $\theta$ is colatitude, see [120] or [23] Eqn. (20)] for details. Fig. 2.2 shows examples of low-degree spherical harmonics.

The naming convention for the spherical harmonic solutions here is analogous to the naming of eigenstates in quantum mechanics: each pressure or gravity oscillation is identified by the triplet of integers $(l, m, n)$, where $l \geq 0$ is the number of node circles on the sphere or spherical harmonic degree, $|m| \leq l$ is the number of node circles through the poles, or azimuthal order, and $n$ is radial order, or number of radial node shells (layers of a spherical onion). The sign of $n$, by convention, indicates the nature of the mode. Pressure modes, or $p-$modes\(^2\) are denoted by $n > 0$, fundamental modes, or $f-$modes have $n = 0$, and gravity modes, $g-$modes have $n < 0$. $f-$modes and

\(^2\)Cowling [28] is responsible for both the mathematical prediction and naming of the $p-$modes.
2.2. ELEMENTARY HELIOSEISMOLOGY

$g-$modes both have gravity as the restoring force, but $f-$modes refer only to surface gravity waves. In terms of frequency the $g-$mode band extends from $10 - 300 \mu$Hz, the $p-$mode band $250 - 5100 \mu$Hz, and the pseudomode band (high frequency $p-$modes) above $5100 \mu$Hz.

Fig. 2.2(b) shows the relation between “quantum numbers” $l$ and $m$ for these solutions. The top plots show node circles, while bottom plots show the appearance of the corresponding surface wave. Pressure modes, when $l$ is fixed and $n$ increases, are approximately equally spaced in frequency, while gravity modes are equally spaced in period. Since there is no preferred axis of rotation when the Sun is not rotating, a dependency on $m$ is introduced when an axis of rotation is actually chosen.

The $p-$ and $g-$mode oscillations are sensitive indicators of structure within the Sun. Acoustic modes, or $p$ modes, have frequencies sensitive to sound speed $c$ in the interior, which is a function of depth $r$. Fig. 2 in [104] shows how pressure modes propagating towards the core are refracted away from the radial direction and reflected back to the surface. Modes with lower $l$ penetrate deeper into the solar interior, and thus their frequencies can be used to deduce properties of the solar interior up to (potentially) > 90% of the solar radius, [37] Fig. 1. While $g-$mode structure originates by nature in the core$^3$, reproducible $g-$mode frequencies are most sought for information about the solar core. Fundamental mode frequencies, or surface gravity modes can be used to make accurate measurements of the solar radius. The inversion problem is the problem of inferring parameters, e.g. sound speed or inner solar rotation velocity, of the solar interior as a function of the mode frequencies and their splittings. Consult [23] for an overview of the various inversion

$^3$Though there is no uncontested experimental evidence for solar $g-$modes to date [7, 5], for various reasons [2, Ch 1], these structures are readily discernible in data from other stars. In the solar case, reproducibility is key: the discovery would be of epochal importance.
2.2. ELEMENTARY HELIOSEISMOLOGY

Asymptotic $p$–mode frequencies can be approximated by the following relation

$$\nu_{l,m,n} \approx \nu_r \cdot m + \nu_0 \left[ \frac{1}{2} \left( l + \frac{1}{2} \right) + n + \alpha \right] + \ldots \tag{2.3}$$

where $\nu_r = 440 \mu$Hz and the rotational splitting term $\nu_0 \approx 135$ nHz, but is dependent upon the mode \[119\]. Also $\alpha$, here taken to be constant, represents the phase change at the point of reflection, see \[23\] p. 11 & Eqn. (47)]. The solar $p$–modes are split into $2l+1$ singlets, indexed by $m$. This asymptotic formula is only a guide as $p$–mode frequencies shift with solar activity \[129\]. Note that the asymptotic formula indicates that $\nu_{l,m,n} \approx \nu_{l+2,m,n-1}$, so these may be difficult to distinguish. For individual modes, the average amplitude is approximately $10$ cm/s \[23\].

Two restoring forces act on the solar modes: pressure and buoyancy (or gravity), which can be described in terms of two important frequencies defined in e.g. \[24\], namely the acoustic frequency $S_l$ and the Brunt-Väisälä frequency $N$. These two frequencies define the regions in the solar interior in which the modes can resonate. $S_l$ depends on $l$ and $r$ and is the frequency of an acoustic wave of degree $l$ traveling horizontally, while $N$ pertains to the frequencies of oscillations in a gas under the effect of buoyancy and depends on pressure, density and radius. The oscillation equations predict that $p$–modes oscillate as a function of $r$ when their frequencies are greater than $S_l$ and $N$, while $g$–modes have frequencies less than $S_l$ and $N$. The region in frequency between these two quantities is evanescent, i.e. the modes are unable to propagate, see \[105\] Fig. 5.12]. Near the surface of the star, modes are reflected where their wavelength exceeds a quantity called the density scale height. The consequence of this theory is that pressure modes with frequencies exceeding the acoustic cutoff
2.3. WHY HELIOSEISMOLOGY?

Clearly, the Sun greatly influences the Earth’s environment as the center of our solar system and the main driver for the Earth’s climate processes. Through its output of charged particles and radiation, the Sun also influences the Earth’s magnetic field and helioseismology has contributed much to our understanding of the solar dynamo, sunspots, and flare activity. For example, through observations of the solar modes, it was inferred that the solar tachocline, the boundary between convective and radiative zones, lay at approximately 0.7 solar radii, much deeper than was originally thought.

The conditions of the interior of the Sun cannot be realized on Earth and thus present a unique view for studying certain fundamental physical processes e.g. the pathways for neutrino production in the Sun’s core, Eqn. (2.1).
As the only star that can be observed in detail, knowledge of the evolution and composition of the Sun is useful for determining those of other, similar stars. For example, stars such as $\gamma$ Dor, $\delta$ Scuti, and slowly pulsating B stars, as well as stars like $\alpha$ Cen exhibit similar oscillations, and thus inversion methods developed from observations of the solar modes is applicable similarly to these stars [109, 24]. The future of helioseismology may lie in the investigation of local helioseismic techniques such as tomography, in order to study subsurface flows near sunspots and highly active regions.

### 2.4 Literature Review

For the sake of completeness, I will give a very brief literature review, which summarizes the salient helioseismic discoveries of the 20th century. A more thorough treatment can be found in either of the excellent review papers [23, 105].

#### 2.4.1 The Discovery of the Solar Modes

The normal modes of the Sun were first discovered in 1962 by Leighton, see [65, 38], as approximately 5-minute oscillations in Doppler measurements of Fraunhofer absorption lines. Remarkably, these authors immediately recognized the potential to use the observed periods to infer properties of the solar atmosphere, in particular the “granulation” which appears as a result of convective flows, see [105, Fig 3.5]. Evidence that the oscillations were standing acoustic waves (proposed by Ulrich, [124]) was corroborated by Deubner et al., [33] who found ridges in the wavenumber-frequency diagram (see §3.1) showing that the oscillations were organized as discrete modes.
In 1976, Hill et al. [56] found that these oscillations produced changes in the apparent solar diameter, which led to the conclusion that these oscillations were global phenomena. This immediately implied that these data would be useful to infer the properties of the solar interior. Claverie et al. [27] showed that modal structures were visible in spatially integrated solar disk images, and that these frequency patterns, for low degrees, matched theoretical asymptotic predictions. Duvall and Harvey [37] observed intermediate degree modes, thereby filling the gap between the results of Claverie for low $l$, and Deubner for high $l$ modes.

The essential contribution of these authors was that the oscillations seen in Doppler velocity measurements in photographic images of the solar disk over time were in the form of discrete modes, oscillating globally. With this came a new wave of *helioseismic inversion techniques* which are used to make detailed inferences of the properties of the solar interior using the modal frequencies and their splittings, e.g. inferences on the internal solar rotation, see Duvall et al. [36], and the sound speed in the interior, see Christensen-Dalsgaard et al. [25].

### 2.5 Modeling of Solar Oscillations

Solar oscillations are thought to be stochastically excited by vigorous subsurface convection with damping [105 §5.4.2]. Loosely speaking, convective motion provides a kind of friction by moving parcels of gas back and forth. This motion additionally provides excitation: imagine a cavity between two reflecting boundaries hit by convectively driven parcels of gas. Such a resonator can be described by a second order differential equation with forcing. Damping in the solar atmosphere is such that waves with frequencies above the acoustic cutoff are confined, see [105 §5.4.2].
I will defer the discussion of modeling of the oscillations until Chapter 5.

Turbulence

Turbulent convection in the superficial layers of the Sun produces unstable vortices and convective granules. The bulk of the kinetic energy in these eddies is contained in the structures operating on large length scales. Energy cascades from large scale structures down to small scale structures produces a superposition of velocity fluctuations by an inertial mechanism. Kolmogorov proposed the statistical theory of turbulence and the notion of energy cascades and self-similarity in 1941 [61], though a complete description of turbulence remains an unsolved problem in physics.

In the present context, the assumed turbulence in the convective zone of the Sun implies large-scale statistical descriptions can be used to model these systems. The overall turbulence relation for the energy spectrum is a power law, which has a dependency in frequency resembling

\[ S(f) = f^{-c}, \quad f \in (0, 1/2) \]  

(2.4)

where \( c > 0 \) is a constant. To identify power law spectra, standard practice is to plot spectrum estimates on a log scale in both frequency and magnitude so that \( \log S(f) = -c \log f \), and the dependence between \( \log S \) and \( \log f \) is linear. Kolmogorov turbulence nominally follows a \( f^{-5/3} \) dependence in data from space and geophysical applications, where the Kolmogorov adjective here implies that the length scales become sufficiently small [43]. Turbulent flows generated in the convection zone imply that solar wind particles inherit the same general form for the energy spectrum. In the data examples given in the next chapter, all spectra shown contain some form
of power law.

2.5.1 Solar Modal Structure in the Interplanetary Magnetic Field and on Earth

It has long been known (see [64] for an interesting historical review) that communications systems, especially wireless communications using satellites and land links, can be adversely affected by energetic particle radiation in the Earth’s magnetosphere. Modern telecommunications satellites must be built to withstand particularly harsh magnetospheric conditions, and the mapping of the radiation environment over time and in space was (and continues to be) a very important component in the design of reliable systems. In particular, the ionizing effects of energizing particles in the interplanetary magnetic field can cause radiation damage of spacecraft electronics and solar arrays, to induced voltages on power cables, and to pipeline corrosion [120].

While investigating the effects of solar activity on communications satellite anomalies and dropped cell phone calls, Bell Labs technical staff Thomson, Lanzerotti and Maclennan reported sharp peaks in interplanetary flux spectra from the Ulysses and Voyager spacecraft whose frequencies corresponded approximately to the normal modes of the Sun, [120]. The implication of this discovery was that the solar wind carried both pressure and gravity mode signatures far beyond 1 A.U., and that these modes strongly permeate the interplanetary magnetic field (IMF)\(^4\). Thomson et al. proposed that the mechanism for the granulation effect near the solar surface (attributed in the standard model to turbulence in the convection zone) could be

\(^4\)The Voyager II spacecraft, launched in 1977, was intended to explore Jupiter and Saturn, and in 1985 was approaching Uranus. It has now (at the time of this writing) entered interstellar space, and will continue to collect data until 2020. The Ulysses mission, launched in 1990 on a pathway to Jupiter, was intended to monitor the sun at high solar latitudes.
2.5. MODELING OF SOLAR OSCILLATIONS

partially the effect of solar gravity waves.

The interplanetary magnetic field is turbulent, and the predominant view at the time was that convection would destroy any trace of modal structure before it exited the solar atmosphere and escaped into the solar wind. Periodogram analyses of IMF data, after all, reported featureless red power law spectra, e.g. [44]. At first Thomson et al. were met with fierce opposition. But, Thomson et al. argued, the appearance of turbulence, while probably correct, does not preclude the coexistence of discrete modes in the IMF, and in the solar wind [118, 43]. Viall et al. independently reported discrete frequency oscillations in the solar wind which coupled with the magnetosphere [126].

The results of Thomson et al. and subsequent papers [118, 119, 43] give evidence that the seismic oscillations of the Sun enter the Earth’s atmosphere where they modify the meteorological conditions, the Earth’s magnetic field and the power grid, and are evident in measurements of the Earth’s seismic state. These findings are of considerable scientific interest, from the reliability of communications systems, to the power grid, to meteorology. The appearance of discrete modal structure in these spectra would imply, in some sense, a more “predictable” energy spectrum than the entirely turbulent theory the solar physics community has become accustomed to. So this particular discovery might prove to be incredibly constructive from the perspective of predictive modeling of adverse solar particle events.
Chapter 3

Data

In this chapter I will introduce some real data and discuss the properties of their spectra.\footnote{The word \textit{spectrum} strictly refers to the spectral density function, see Appendix A, but in this thesis and in the literature, when it is clear from the context, can refer to the spectral density, an estimate of the spectral density, or an estimator of the spectral density. Note that the term \textit{standardized spectrum} refers to a particular normalized spectrum estimator or estimate.} These examples will be used to motivate the theoretical contributions of this thesis. I will begin with optical Doppler measurements of the solar disk, and then move outward from the Sun to the first Lagrange point where the SWEPAM instrument aboard ACE monitors the solar wind, and then to Earth where neutron and muon monitors are observing secondary cosmic radiation. In each of these data sets I will show examples of time series containing modal structures, and demonstrate some of the difficulties encountered when examining data suspected to contain many line components. Note that of the data sources presented in this chapter, only the optical Doppler measurements, and more controversially the solar wind measurements have been documented in the literature to contain these structures. The implications of strong modal structures in cosmic rays observed on Earth challenges our understanding of the extent to which the solar magnetic field distorts the galactic charged...
3.1 Optical Doppler measurements

A natural place to begin investigations of helioseismic spectra is the data with which the solar modes were discovered - optical Doppler velocity measurements. The GONG network, [52], comprises a set of 6 terrestrial observatories which have been continuously producing $246 \times 242$-pixel solar disk images with one minute sampling since October 1995. The goal of the project is to record the Doppler velocity of the solar surface for the purpose of determining solar $p$- and $f$-mode frequencies and rotational splittings. The spherical harmonic functions (2.2) are used as a basis for an orthogonal expansion over disk images of the visible portion of the Sun’s surface over time, resulting in a time series of amplitudes specific to each pair $(l,m)$.

The diagnostic diagram can be computed by averaging the power spectra of these time series over all $m$ and placing these side by side which results in a 3D plot - a spectrum which evolves with $l$. Figure 3.1 shows such a diagram calculated from 36 days of 1-minute GONG data beginning on July 10, 2004 for $0 \leq l \leq 30$, see also [104] Fig. 3 or [52] Fig 1C. The diagnostic diagram shows that the power distribution of these oscillations, in terms of degree $l$ and frequency $\nu$, has peaks in discrete ridges, that is, the oscillations are discrete and correspond to modes with an increasing number of radial nodes ($n$ increasing). The diagnostic diagram shows a lot of information at once about the nature of the experimentally-observed solar modes.

The $p$-mode frequencies are determined from spectra such as that shown in the left panel of Fig. 3.1 by fitting a Lorentzian model for the harmonic oscillations. These methods are discussed later on in the thesis.
3.1. OPTICAL DOPPLER MEASUREMENTS

Figure 3.1: Panel (a) shows a spectrum of the 1-minute GONG data for $l = 0$, median spectrum taken over $m$. Panel (b) gives a diagnostic diagram computed for $0 \leq l \leq 30$. The mean spectrum was taken of the top 2/3rds of the estimates for each $l$ over the time series indexed by $|m|$. A combined Welch-multitaper spectrum was used where multitaper parameters are $NW = 4, K = 6$ and the data blocks were one day long. The distinct ‘ridges’ visible in this diagram can be indexed in $n$, the number of radial nodes. The colour map has arbitrary units in power.

One has to be careful about leakage of power from neighboring modes when performing parametric model fitting on these spectra. Though the spherical harmonic basis is orthogonal over the entire surface of the Sun, the images capture only the part visible from Earth, and curvature near the edges of the images means that only about 1/3 of the solar disk is visible at any given time, so the spherical harmonic basis is not leak-proof. Furthermore, the oscillations are highly temporally variable.
3.2. SOLAR WIND DATA

Stix [105, p. 183] notes that “one–third oscillates”, that is, one third of the Sun is oscillating at any given time or location on the Sun which implies that one should be careful in using robust estimates for such spectra. In this case the use of robust estimates might result in exclusion of the solar modes altogether. The left panel of Fig. 3.1 shows the $l = 0$ spectrum from this 36-day window estimated using a Welch technique retaining only the largest 2/3 of the spectra.

3.2 Solar Wind Data

Solar wind data contain many lines. Data collected from the SWEPAM instrument aboard the Advanced Composition Explorer (ACE)\textsuperscript{2} is used here to demonstrate the statistical properties of these spectra. Note that these data have been analyzed in previous studies, e.g. [119] and [43] which both show spectra with frequencies below 500$\mu$Hz, i.e., near the beginning of the $p$–mode band. Here, higher frequencies $\sim 2000 – 3000\mu$Hz, i.e., mid-range $p$–mode band are analyzed.

Beginning with data from January 1, 1999 to September 24, 2013 with a 64-second sampling rate, (which yield spectra with Rayleigh resolution of 2.15 nHz and 7.8mHz Nyquist frequency) the left panel of Fig. 3.2 shows raw data. Note that the proton density measurements are nonnegative, which implies a non-Gaussian distribution. Mean, variance, skewness and kurtosis for these data are $\bar{x} = 6.17$ protons/cm$^3$, $s^2 = 28.82$, $g_1 = 4.06$, $g_2 = 33.40$. Interpolating missing data and computing an initial multitaper spectrum estimate with $NW = 4, K = 6$, seen in the right panel of Fig. 3.2 yields a spectrum which appears to contain a power law, except for some peaks in the extreme low-frequency range.

\textsuperscript{2}These data are publicly available from \url{http://www.srl.caltech.edu/ACE/ASC/level2/lvl2DATA_SWEPAM.html}
The translated $f^{-5/3}$ characteristic of Kolmogorov turbulence may be approximately correct for the ACE data, as a linear least-squares line fit to the log-log spectrum between $f = 10^{-5}$ and $10^{-2}$Hz has slope $-1.689^3$.

Figure 3.2: Proton density data from the ACE SWEPAM instrument between Jan 1, 1999 and September 24, 2013. Panel (a) shows the data, while panel (b) shows the multitaper spectrum ($NW = 4, K = 6$) The dashed line is a (shifted) $f^{-5/3}$ power law characteristic of Kolmogorov turbulence.

Standardizing the spectrum on the band $2000 – 3000\mu$Hz band, mid $p$–mode band, by removing a linear trend as in the left panel of Fig. 3.3 gives the spectrum seen in the right panel of Fig. 3.3 The reason for not extending this band to the entire $p$–mode band was because significant, unexplained, transient harmonics$^4$ are present near $4000\mu$Hz, which may be real, or could reflect some instrumentation noise.

$^3$Because the spectrum estimate at adjacent frequency bins is non-negatively correlated, the assumption of independent samples appropriate for linear regression is violated and it would be misleading to quote $R^2$.

$^4$These artifacts were first noticed by Queen’s Natural Sciences Engineering and Research Council (NSERC) Undergraduate Student Research Assistant (USRA) Julian Fortin.
3.2. SOLAR WIND DATA

Figure 3.3: Panel (a) shows a linear trend which is removed from the log-log spectrum so that harmonic components can be separated from the power law background. The resulting standardized spectrum from the proton density data is shown in (b). Dashed lines denote significance levels for the standardized spectrum. The spectrum pictured has $\approx 2K = 12$ degrees of freedom (df). There are a large number of peaks above the 99.9999% significance level.

Distribution of ACE spectrum

The distribution across frequency of this seemingly “largely modal” spectrum, begs some investigation. In particular, given an estimate of the spectrum with $\nu$ degrees of freedom, an estimate of the fraction $\epsilon$ of the frequency bins having line components is desired, as well as an estimate of the signal to noise ratio $\rho$, see Appendix A for a discussion of the distribution of spectrum estimates. In [118], minimum distance estimation is applied to spectrum estimates in order to estimate the desired
3.2. SOLAR WIND DATA

parameters.\textsuperscript{5}

Figure 3.4: The results of a simulation of the mixture model described in the text. For this test, \( J = 100 \) is used with \( \nu = 12 \) \( \epsilon = 0.4 \) and \( \rho = 1.2\nu \). Panel (a) shows that the minimum of the \( D \) statistic occurs when \( \hat{\epsilon} = 0.36 \) and \( \hat{\rho} = 1.1\nu \), estimates which are reasonably close to the true value. Panel (b) shows a histogram of the simulated data (solid line) and the best-fitting mixture (dashed curve) of central (dot-dashed curve) and noncentral (dotted curve) chi square distributions in dashed lines.

To quantify the proportion of a spectrum estimate in modes, it is helpful to decompose the distribution of \( y(f) = 2\alpha z(f) \) across frequency into central \( \chi^2_{\nu} \) and noncentral \( \chi^2_{\nu}(\mu) \) components, where \( \mu \) denotes the noncentrality parameter, and \( \nu = 2\alpha \). The idea is that the spectrum has a mixture distribution of central and noncentral \( \chi^2 \) components, the noncentral portion coming from line components in

\textsuperscript{5}Minimum distance estimators \cite{128, 81} are obtained by minimizing a notion of distance between the empirical distribution function of some random quantity and some parametric family of distribution functions. The advantage of the minimum distance technique over other possible parameter estimation techniques (such as maximum likelihood (ML), method of moments (MM), etc. see \cite{81}) is that when conjectured parametric model is incorrect, the minimum distance estimator is typically consistent with respect to the particular distance metric used, that is, the selected model will always select the best approximation in the class of parametric models which lies closest to the empirical distribution with respect to the particular distance metric chosen. The case for the noncentral chi-square distribution in which the df parameter is known is considered by Anderson \cite{3}, and it is the approach presented there and in \cite{118} that I follow here.
the spectrum. Decomposing the distribution of the $\chi^2$ standardized spectrum should give an estimate of the amount of “contamination” coming from the modes. Modifying slightly so that $\mu$ is in units of $\nu$, let $\rho = \mu/\nu$, and this quantity will denote the signal to noise ratio (SNR) when the modes are the signal and the background (approximately white) spectrum is the noise. Proceeding in the spirit of [118], assume that the modal frequencies all have a common noncentrality parameter, and write the cdf of the spectrum as an $\epsilon$-contaminated mixture distribution

$$F(y; \epsilon, \nu, \rho) = \epsilon P_c(y; \nu) + (1 - \epsilon) P_{nc}(y; \nu, \rho)$$

(3.1)

where $P_c$ denotes the central, and $P_{nc}$ the noncentral parts of the distribution with proportions $\epsilon$ and $(1 - \epsilon)$, respectively, for $0 \leq \epsilon \leq 1$. Let the empirical distribution of $y(f)$ be denoted by $\hat{F}(y)$.

Then the objective is to estimate $\epsilon, \rho$ so that the difference between the empirical distribution and the theoretical distribution is as small as possible. The approach presented in [118] is to minimize a goodness-of-fit test statistic, namely the Kolmogorov-Smirnov distance

$$D(\epsilon, \rho) = \max_y |F(y; \epsilon, \nu, \rho) - \hat{F}(y)|$$

(3.2)

by choosing $\epsilon, \rho$ on a closely spaced grid of frequencies, and then choosing the pair $\epsilon_{\text{min}}, \rho_{\text{min}}$ which minimize $D$ over all choices of $\epsilon, \rho$. $\hat{F}(y)$ denotes the empirical distribution of the spectrum. $D$ is a function of $\epsilon, \rho$ alone because in practice the degrees of freedom are approximately $2K$. In practice, one scales the spectrum so that the mean of the spectrum matches the theoretical mean of the distribution [3.1].
which is achieved by multiplying the sorted samples from the spectrum $s(j)$ by the factor $\gamma = \frac{\nu + \rho}{(1/J) \sum_j s_j}$. This method is demonstrated on simulated data in Fig. 3.4 by generating $J = 100$ samples of which 40 are noncentral $\chi^2_{12}(14.4)$ and 60 are central $\chi^2_{12}$. The contour plot shows the choices of $\epsilon, \rho$ for which the Kolmogorov-Smirnov $D$ statistic is computed. The minimum value of $D$ occurs when $\epsilon = 0.36$ and $\rho = 1.1 \nu = 13.2$ which is fairly close to the original values set for the simulation.

Figure 3.5: The results of the decomposition of the scaled ACE standardized spectrum, $\nu = 12$, in Fig. 3.3(a) into noncentral and central components is shown. Panel (a) is a contour diagram of the minimization of the $D$-statistic for a grid of choices for $\epsilon, \rho$. The minimum of $D_{\text{min}} = 0.0035$ occurs when $\hat{\rho} = 0.6\nu = 7.2$ (labeled $\lambda$ above) and $\hat{\epsilon} = 0.46$. Note that the best fitting central distribution has $D_{\text{cent}} = 0.0462$. The histogram in (b) illustrates the fit of the best fitting mixture distribution. The number of frequency bins included in the analysis was 2,576,980 samples or 464,896 Rayleigh resolutions.

When one decomposes the empirical distribution of the ACE data, one obtains the result in Fig. 3.5 namely that about 46% of the frequency bins are modal, and the estimated noncentrality parameter is $0.6(12) = 7.2$, indicating a very large number of lines with a fairly low signal-to-noise ratio. This implies that $\sim 30\%$ of the total power
is in the noncentral component. Note that the $D$ statistic for the central distribution is 0.0462. This analysis gives reason to believe that these data contain many line components.

These results are reasonably consistent with those in [118] in which solar flux data from Ulysses are decomposed on the $g$–mode band between $60 – 80\mu$Hz. The mixing fraction $\epsilon \approx 0.36$ and SNR $\rho/\nu = 0.8$ were estimated to give a total of 30% of the total power in the noncentral component.

3.3 Cosmic Ray Data

*Galactic cosmic rays* are high energy charged particles entering the atmosphere from outer space. For an interesting historical review see [21]. Hess’s 1912 [55] discovery of cosmic radiation used electroscopes aboard a hydrogen filled balloon (reaching an altitude of 500m) to conclude that the ionization of the atmosphere increased with altitude. Wulf [131] had found, after a famous 1909 experiment conducted from the Eiffel Tower using an improved electroscope, that either the absorption length for gamma rays in air was bigger than the prevailing estimate or there must be another source of atmospheric radioactivity. Convinced that this radiation was electromagnetic, Millikan gave this ionizing radiation the term “cosmic ray”, though the composition of the radiation was poorly understood at the time [74], and it was later resolved by experiments of Clay, Compton, and Millikan, around 1933 [75], that the ionizing radiation had also a corpuscular\(^6\) component.

In fact, cosmic radiation is best divided into primary radiation and secondary radiation. The primary radiation originates from various sources such as star deaths,

\(^6\)Corpuscular radiation refers to radiation of subatomic particles such as electrons, protons, neutrons, or alpha particles.
quasars, neighboring suns, etc., while the secondary radiation is caused by the collisions of the primary radiation with atomic nuclei in the Earth’s atmosphere. The primary radiation consists mainly of charged particles and gamma ray photons. The secondary radiation is diverse and can be divided into three broad categories (i) the nucleon component, which consists of protons and neutrons (ii) the meson component, of muons and (iii) the soft component, of electrons and gamma ray photons [93]. I will present data collected by terrestrial neutron monitors in this thesis.

Briefly, the study of cosmic radiation is important because it gives insight into the meteorology of outer space and the origins of the universe. Knowledge of the ionization properties and variation of cosmic ray intensity over time has led us to better understand, for example, the mechanism for ozone depletion in the atmosphere and the radiation environment for space physics experiments in outer space and satellites orbiting the Earth. I have already shown in Fig. 2.1 that when solar activity is high, secondary cosmic ray incidence on Earth is low, and in this way we perceive that the observation of cosmic radiation offers an interesting view of the magnetic activity of the Sun.

3.3.1 Neutron Data

A neutron monitor is a shielded device which counts the number of high energy neutrons arriving at its core. Neutrons are convenient particles for study since they are minimally attenuated while moving in a charged medium, though neutron behavior in the atmosphere changes with latitude and altitude and is sensitive to local atmospheric pressure and meteorological conditions.

Neutron monitors consist of an array (2m² or larger) of proportional counters
3.3. COSMIC RAY DATA

filled with boron trifluoride (BF$_3$). When neutrons enter the device, they collide with the lead nuclei and are deflected in all directions. Collision with proton nuclei in the paraffin which coats the counters brings the neutrons into the thermal range (about 0.025eV), whereupon the thermal neutrons will ionize B$^{10}$ nuclei in the boron trifluoride. The inner paraffin layer is rich in protons and helps direct neutrons which have reached the counters. This process translates each thermal neutron ionization event into a pulse of electricity. The resulting data is given as multiplicity, or counts per unit time.

Real-time neutron monitoring is useful for (i) determining when cosmic ray particles are a result of rare solar events, such as Forbush decrease and ground level enhancement events, (ii) calculation of the ionization of the atmosphere (iii) information about anisotropy and pitch angle of incident cosmic rays help to determine certain conditions of interplanetary space [71].

Data Analysis

Neutron monitor data was provided by the Bartol Research Institute, from Thule, the South Pole, and Newark in 10-s intervals between June 30, 2004 and Dec 31, 2005 and is shown along with concurrent atmospheric pressure data in Fig. 3.6. The South Pole dataset is somewhat shorter, ending on October 14th, 2005, when the instrument was taken off line and was remounted some distance away. Ten second data was requested for the purpose of investigating neutron data for modes even though the $p$–mode band nominally ends at about 5100µHz as a result of the aforementioned acoustic cutoff. The main reason for this is to avoid aliasing, but secondary to this is to check if the acoustic cutoff applies to the interplanetary magnetic field (IMF) and cosmic rays.
3.3. COSMIC RAY DATA

Figure 3.6: Top plots show raw neutron monitor multiplicity (counts/10s) from Thule, Newark, and South Pole from June 30, 2005 to December 31, 2005 for Newark and Thule and from June 30, 2005 to October 14, 2005 for the South pole. Data have not been interpolated and outlying data points have not been removed. Bottom plots show atmospheric pressure (in mmHg) at each site over time. Note differences in the scales in the y-axis.

The Bartol Research Institute also has quality 1-minute and 1-hour data going back as far as 1957 in Thule.

The anomaly seen in Fig. 3.6 on 21-22 January 2005 appears at first to be an outlier, but persists for more than just a few samples and is consistent throughout the three datasets. This event corresponds to a well-documented geomagnetic storm, see e.g. [34]. Incidentally, from Fig. 2.1 one perceives that the solar magnetic cycle (using the sunspot number as a proxy) is heading into a minimum during this period, which is preferable since galactic cosmic rays are best detected when solar magnetic activity is at its quietest.

Fig. 3.6 shows the relationship between neutron incidence and atmospheric pressure. Neutron attenuation by local atmospheric conditions is described by a negative exponential relationship. When atmospheric pressure is high neutron multiplicity
drops accordingly. In order to remove this dependency, neutron data are often “pressure corrected”, the details of which can be found in e.g. [67]. The objective of the atmospheric pressure correction is to remove the dependency on atmospheric conditions, and results in a signal which (visually) appears to be stationary.

Figure 3.7: Neutron monitor spectra (low frequency) from pressure-corrected Thule, Newark, and South Pole data. Data are plotted on a log-log scale for comparison. The Newark neutron monitor data has three strong sinusoidal components corresponding to harmonics of 11.57 µHz, or the daily cycle. Power laws fitted to the log-log spectra on the domains indicated have powers $-1.34$, $-1.44$, and $-2.02$ for Thule, Newark, and the South Pole, respectively (here the lines are shown a decade below the corresponding least squares fit). Note the smaller overall power of the Newark spectrum, which indicates smaller variance. Multitaper parameters $NW = 5$ and $K = 8$ were used. The same time intervals shown in Fig. 3.6 were used.

Summary Statistics

Summary statistics for the raw and pressure-corrected data are reported in Table 3.1 below, with obvious outliers and the geomagnetic storm of Jan/2005 excluded.

Neutron monitors have an error inherent in the detection process, that is, occasionally two thermal neutrons originating from the same collision will take different amounts of time to trigger ionization in the proportional counter. This will result in
3.3. COSMIC RAY DATA

<table>
<thead>
<tr>
<th>Raw Data</th>
<th>Thule</th>
<th>Newark</th>
<th>South Pole</th>
</tr>
</thead>
<tbody>
<tr>
<td>Number of 10-s samples, $N$</td>
<td>4,751,995</td>
<td>4,751,999</td>
<td>4,412,997</td>
</tr>
<tr>
<td>Rayleigh resolution, $R$</td>
<td>21.04nHz</td>
<td>21.04nHz</td>
<td>22.66nHz</td>
</tr>
<tr>
<td>Data missing (%)</td>
<td>0.9</td>
<td>16.6</td>
<td>1.5</td>
</tr>
<tr>
<td>Sample mean, $\bar{x}$ (counts/10s)</td>
<td>205.3</td>
<td>96.2</td>
<td>328.2</td>
</tr>
<tr>
<td>Variance, $s^2$</td>
<td>349.3</td>
<td>70.0</td>
<td>759.2</td>
</tr>
<tr>
<td>Skewness, $G_1$</td>
<td>0.062</td>
<td>-0.326</td>
<td>-0.046</td>
</tr>
<tr>
<td>Kurtosis, $G_2$</td>
<td>3.19</td>
<td>3.93</td>
<td>2.60</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pressure Corrected Data</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean, $\bar{x}$</td>
<td>190.3</td>
<td>93.4</td>
<td>325.0</td>
</tr>
<tr>
<td>Variance, $s^2$</td>
<td>187.4</td>
<td>25.8</td>
<td>180.2</td>
</tr>
<tr>
<td>Skewness, $G_1$</td>
<td>0.136</td>
<td>-0.271</td>
<td>-0.073</td>
</tr>
<tr>
<td>Kurtosis, $G_2$</td>
<td>3.24</td>
<td>4.06</td>
<td>2.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Atmospheric Pressure Data</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sample mean, $\bar{x}$ (mmHg)</td>
<td>753.9</td>
<td>758.7</td>
<td>509.4</td>
</tr>
<tr>
<td>Variance, $s^2$</td>
<td>58.8</td>
<td>35.5</td>
<td>50.4</td>
</tr>
<tr>
<td>Skewness, $G_1$</td>
<td>0.062</td>
<td>-0.322</td>
<td>-0.105</td>
</tr>
<tr>
<td>Kurtosis, $G_2$</td>
<td>3.19</td>
<td>3.93</td>
<td>2.77</td>
</tr>
</tbody>
</table>

Table 3.1: Summary statistics for neutron data from the three sites. These data may be interpreted as reasonably Gaussian, but do not have equal mean and variance, as one would expect from a Poisson counting process. For example, raw data from Thule has sample mean 205.3 counts/10s while the variance, 349.3 (counts/10s)$^2$, is $\sim 75\%$ larger. Skewness and kurtosis of Thule raw and atmospheric pressure data disagree in the fourth decimal place.

two detections occurring instead of just one. This manifests itself as higher variance in the resultant data than is expected of a counting process, that is, overdispersion, and often it is necessary to adjust the variance by a multiplicity factor, which varies from monitor to monitor.

Neutron monitor data are reported as counts per 10-s period, thus the process ought to be roughly Poisson in nature, i.e. $\bar{x} \approx s^2$, but appears to be underdispersed in the case of the data from Newark and pressure corrected data from the South Pole but overdispersed in the case of raw data from the South Pole. The data from
Thule, for example, appears to have been corrected so that its first two moments are approximately equal, while the sample skewness has increased slightly.

Spectrum Analysis

<table>
<thead>
<tr>
<th>mode</th>
<th>freq (µHz)</th>
<th>p(3,13)</th>
<th>p(1,14)</th>
<th>p(2,14)</th>
<th>p(0,15)</th>
<th>p(3,14)</th>
<th>p(1,15)</th>
</tr>
</thead>
<tbody>
<tr>
<td>South Pole</td>
<td>2138.05</td>
<td>2157.08</td>
<td>2217.96</td>
<td>2228.52</td>
<td>2273.55</td>
<td>2292.14</td>
<td></td>
</tr>
<tr>
<td>Thule</td>
<td>2137.91</td>
<td>2156.76</td>
<td>2217.50</td>
<td>*</td>
<td>2273.49</td>
<td>2292.02</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.8: Standardized South Pole (left panel) and Thule (right panel) pressure corrected neutron monitor spectra, 2130-2300 µHz band shown. Horizontal lines denote significance levels for a $\chi^2_{16}$ standardized estimate. Vertical lines denote $p$—mode frequencies in that band for $l \leq 3$ computed from BiSON data. All of the mode frequencies correspond to peaks in this spectrum, for Thule 2/6 have significance above 99.9% and 5/6 have significance above 99%, for the South Pole, 2/6 have significance above 99%. The table below shows the mode names and frequencies (µHz), with the locations at which the mode peaks occur in the two spectra. In these examples, the significant peaks lie within $\sim 1R$ and are marked with * if there is no peak above 90% within $1R$. For example, a line corresponding to a $p$—mode frequency at 2,156.8µHz with $l = 1$ and $n = 14$ is present at better than 99.9% significance in both spectra.

Neutron monitor data were sampled from June 30, 2005 to December 31st, 2005,
interpolated, low-pass filtered and decimated to $\delta t = 90s$. The Rayleigh resolution of the data series are shown in Table 3.1. Because fine scale frequency resolution is desired, the time bandwidth $NW = 5$ was chosen with $K = 8$ tapers, giving a bandwidth of about $\pm 105\text{nHz}$ for Newark and Thule and $\pm 113\text{nHz}$ for the South Pole spectra. Recall that asymptotically, $p$-modes (2.3) are split by approximately $440\text{nHz}$, so this choice of bandwidth should (in theory) allow resolution of both the modes and their splittings, in the rare case where interference with other modes and/or noise does not obscure the pattern of $2l + 1$ equally spaced singlets. Also, the singlets ($\pm m$, or sectoral modes) often have considerably more power than the center ($m = 0$) frequency so detectability is uneven. The choice of bandwidth was also motivated by the need for additional degrees of freedom in multitaper estimates when the range of the spectrum is large (as is the case with power law spectra). To reduce the range of the spectrum to maximize the adaptive weights, an AR-1 model was fitted and removed from the data.

Figure 3.8 shows a standardized spectrum estimate of pressure corrected neutron monitoring data obtained from the Thule neutron monitor. Superimposed on this spectrum are $l \leq 3$ p-mode frequencies obtained from Birmingham Solar Oscillations Network (BiSON) [20] (vertical lines in Fig. 3.8), labeled in the accompanying table. There is agreement between the 6 frequencies pictured on this band, most having peaks with better than 99% significance in the spectrum. Note that the standardization is for a Gamma(16, 1/16) process, and that significance levels (horizontal lines in the figure) are labeled under the assumption that no modal structure is present.

The probability density function (pdf) for the standardized spectrum is Gamma $(\alpha, 1/\alpha)$ distributed, where $p(z, \alpha) = \alpha^\alpha z^{\alpha-1} \exp(-\alpha z)/\Gamma(\alpha)$, where $\Gamma(\cdot)$ is the gamma function, see Appendix A.
Abundant modal structure causes this spectrum to appear pulled upward; the benchmark level used to standardize the spectrum is the 5% point.

Figure 3.9: Decomposition of the Thule, Newark and South Pole neutron spectra on the 2-3mHz band. Top plots show contours of the Kolmogorov-Smirnov D statistic (see text) and the bottom plots show the histograms of each estimate with the best mixture of central and non central $\chi^2$ distributions plotted in red on top. Green and blue dotted curves represent $(1-\hat{\epsilon})$ scaled central and $\hat{\epsilon}$ scaled noncentral distributions, respectively.

Analyzing the distribution of the multitaper spectrum estimates, I have decomposed each spectrum in the middle of the $p$–mode band (2000 – 3000)$\mu$Hz as a mixture model of central (“noise”) and noncentral $\chi^2$ (“lines”) components, finding the mixing fraction $\epsilon$ and noncentrality parameter $\rho$ by minimizing the K-S goodness-of-fit test statistic $D$ on a fine grid. Table 3.3.1 summarizes these findings,
3.3. COSMIC RAY DATA

<table>
<thead>
<tr>
<th></th>
<th>Thule</th>
<th>Newark</th>
<th>South Pole</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \dot{\epsilon} )</td>
<td>0.56</td>
<td>0.46</td>
<td>0.60</td>
</tr>
<tr>
<td>( \dot{\rho} )</td>
<td>0.50( \nu )</td>
<td>0.45( \nu )</td>
<td>0.45( \nu )</td>
</tr>
<tr>
<td>% by power ( (\dot{\epsilon}\dot{\rho}/\nu) )</td>
<td>29.1</td>
<td>20.7</td>
<td>27.0</td>
</tr>
<tr>
<td>( D_{central} )</td>
<td>0.0429</td>
<td>0.0392</td>
<td>0.0304</td>
</tr>
<tr>
<td>( p_{central} )</td>
<td>0.0018</td>
<td>0.0058</td>
<td>0.0596</td>
</tr>
<tr>
<td>( D_{min} )</td>
<td>0.0053</td>
<td>0.0037</td>
<td>0.0022</td>
</tr>
<tr>
<td>( p_{min} )</td>
<td>( 1 - 1 \times 10^{-9} )</td>
<td>( &gt; 1 - 1 \times 10^{-15} )</td>
<td>( 1 - 3 \times 10^{-8} )</td>
</tr>
</tbody>
</table>

Table 3.2: Empirical distributions of Thule, Newark and South Pole spectra are decomposed into a mixture of central and noncentral components on 2-3mHz and the resulting parameters are reported here. Note the relatively small values of the SNR (0.5, 0.45 and 0.45 for Thule, Newark and South Pole, respectively). \( D_{central} \) is the value of the K-S D statistic when \( \dot{\epsilon} \) is set to 0, with \( p \)-values.

while Fig. 3.9 graphically shows the minimization of the \( D \)-statistic over the grid. Here \( \nu = 16 \), degrees-of-freedom for the multitaper spectrum (using \( NW = 5, K = 8 \)). \( \rho/\nu \) is the signal-to-noise ratio, and \( D_{central} \) and \( D_{min} \) refer to the value of the K-S test for the central distribution and at the minimum. South Pole and Thule spectra are convincingly 56% and 60% modal, while Newark is \( \sim 46\% \) modal. \( p \)-values are assigned according to [12, 102] (the large sample formula is used).

Splittings

Resolving splittings of the \( p \)-modes is exceedingly difficult. With a bandwidth of approximately 100nHz as is the case here, visualizing splittings is theoretically possible. Unfortunately because modes are not isolated, but interfere with one another, it can be difficult to isolate examples of modes resonating strongly, and with all expected

\[ \text{Note that both papers contain small typos: [12] contains an error in equation (1.1), the term } z/N \text{ should read } z/\sqrt{N}, \text{ while [102] has an } e^{-r^2z^2} \text{ term which should instead read } e^{-2r^2z^2}. \]
3.4. CONCLUSIONS

Figure 3.10: **Leftmost panel:** Thule pressure corrected spectrum estimate, BiSON \( p - (3, 14) \) mode frequency centered at \( 2273.563 \pm 0.026 \mu \text{Hz} \), splittings are assumed to be \( 440n\mu \text{Hz} \) apart. The center frequency corresponds to a peak with more than 99.99 significance. **Middle panel:** Newark pressure corrected data, MDI \( p - (6, 12) \) mode frequency centered at \( 2142.5515 \pm 0.0098 \mu \text{Hz} \); many of the splittings appear with over 99% significance. **Right panel:** South Pole pressure corrected data, MDI \( p - (7, 11) \) center frequency at \( 2045.9401 \pm 0.0055 \mu \text{Hz} \); the center frequency appears with high significance, but splittings are not present with high significance. There may be a slight shift in frequency over time, as seen by the peak shape. In all, if one counts peaks as in communication theory, 1 being “peak is present above 50% level” 0 as “spectrum is below 50%”, one gets 7/7, 10/13 and 9/15 correct for the three examples above.

Singlets present. Between about 2000 and 5000 \( \mu \text{Hz} \), \( p \)-modes have an average spacing of about 10\( n\mu \text{Hz} \), which means that over three years of data would be required to resolve just the average spacing \[119\]. Fig. 3.10 shows examples from all three datasets. Keep in mind that modal singlets may appear with variable power.

3.4 Conclusions

In all of the examples from this chapter, the underlying problem has been the detection or coincidence detection of an exceedingly large number of modes in coloured
noise. When mode frequencies are known, there are a large number of tests for significance of modal components (and their rotational splittings), and their simultaneous assessment becomes a very large multiple comparisons problem. The approach thus far has been either not to adjust for multiple comparisons, or to use a combinatorial approach to get the probability of achieving any particular combination of modes.

Multiple comparisons problems with independent samples are solved in textbooks by, e.g. Bonferroni correction, or less conservative methods of p-value adjustment, see [130] for a summary. For example, controlling rates of false detection (FDR) can be shown to be highly desirable, even when the tests have positive regression dependency [10,11]. In the next chapter I will take the latter approach and determine the expected number of falsely rejected hypotheses, that is, spurious significant peaks in the spectrum. A major difficulty, however, in determining the false discovery rate is that the number of possible hypotheses to be tested is assumed to be known in advance, but this is not practically possible for spectrum estimation. I will have more to say about falsely detecting various combinations of modes at the conclusion of the next chapter.

Another complication with coincidence detection of modal frequencies is frequency modulation (FM). Due to temporal changes in the angular rotation rate of the Sun near the solar tachocline (the rotational shear near the tachocline is thought to generate the solar dynamo), Howe et al. [57] showed, using helioseismic inversion techniques, that near the solar equator, there was a strong periodic ~ 1.3 year dependency, which describes the temporal variations in the subsurface solar rotation rate. Furthermore, this cyclic dependency has been described in the solar wind fluctuations by Szabo et al. [107]. This implies that changes over time of the modal frequencies
and their splittings is occurring.
Chapter 4

Spacing and Shape of Peaks in Nonparametric Spectrum Estimates

When one is analyzing time series which are suspected to contain many lines, the usual approach is to assume for the null hypothesis, $H_0$, that the data do not contain periodic or other deterministic components. Alternatively, $H_1$, the process contains many, perhaps several thousand, weak, approximately periodic components. This pair of hypotheses is appropriate when analyzing composite spectra, or data suspected to contain all of the solar modes. When $H_0$ is the case, the spectrum estimate follows a central $\chi^2$ distribution and the central $\chi^2$ assumption can be used to test for significant peaks in a standardized spectrum. False detections of modal signatures can be numerous, and peaks in spectrum estimates indicating significant lines may appear in patterns indicative of modal splittings. In this Chapter we ask how are the significant excursions of the standardized spectrum bunched: either as (i) a relatively small number of wide peaks above some high significance level, say 99%, or (ii) a relatively large number of very narrow ones? If the spectrum is made up of thousands of genuine line components, as in case $H_1$, the spectrum estimate should contain a
large number of narrow peaks as in (ii). If \( H_0 \) is the case, all perceived lines are spurious, so significant peaks in the estimate should be few - but how many peaks should be expected in this case? In other words, \textit{what is the rate of false detection of lines in a spectrum estimate?}

In this chapter the expected rate of spurious peak detection of the standardized spectrum is given. These results are derived using the formulae of Rice \cite{Rice1989} and others concerning the crossing rate of \( \chi^2 \) distributed processes. As spectrum estimates are \( \chi^2 \) distributed under \( H_0 \), the crossing expression is valid for a prewhitened spectrum (or standardized spectrum, see appendix A), and the expected number of \textit{upcrossings} of a given significance level is obtained. A particularly convenient \textit{standardized} form of the spectrum used throughout this thesis is

\[
z(f) = \frac{\hat{S}(f)}{S(f)}
\]

which is distributed as a scaled central \( \chi^2_{2\alpha} \) random variable, and takes a \text{Gamma}(\alpha, 1/\alpha) probability distribution. Throughout this chapter, \( 2\alpha \) is the number of degrees of freedom of the spectrum estimate and \( C_R = NW \) is the multitaper time bandwidth product. This distribution has mean 1 because \( E\{\hat{S}(f)\} = S(f) \) when the spectrum is constant and the variance is \( 1/\alpha \). Standardized spectra are plotted on a linear scale instead of a log scale in the \( y \)-axis. For a discussion of why it is necessary to standardize the spectrum and how to standardize the spectrum in practice, refer to \[A.3\].

For high levels of significance, an upcrossing of the spectrum is generally associated with only one peak, so spurious upcrossings of a given level in general correspond to spurious peaks in the spectrum. The level-crossing rate of a process is naturally
dependent upon the derivative of the process as well as the distribution of the process itself, specifically the covariance structure of the derivative of the process. Thus the expression for the false detections in a spectrum estimate is dependent upon the variance of the derivative of the spectrum estimator, or simply the autocorrelation function of the spectrum estimator, which is called the antecorrelation [110, 117].

The rate of false detections of a spectrum estimator is also related to the average amount of time the spectrum spends above a given level once that level is exceeded - namely the dwell band, or above-level dwell distance. When analyzing the peaks exceeding the $100(1 - \alpha)\%$ significance level, the difference between wide peaks and narrow peaks should be emphasized, as these have very different implications. I will show in this chapter that spurious peaks in a multitaper spectrum estimate have a triangular shape, and are wider than genuine line components. That is, the shape of the spectrum near a peak, or the autocorrelation of the spectrum estimate near a peak, is indicative of whether or not the peak shows a genuine line component.

In the next section, the results of Rice are reviewed, as well as the extensions to the $\chi^2$ processes. In the following sections, these results are applied to standardized spectra.

### 4.1 Crossing Problems

The Rice equation describes the rate of level crossings of a stationary Gaussian process. Because spectrum estimates are $\chi^2$ distributed, an extension of the Rice equation for $\chi^2$ processes is used and applied directly to the standardized spectrum. Where the Rice equation relies on the spectrum of the process, the $\chi^2$ extension applied to the spectrum relies on the spectrum of the spectrum. The standardized units for the
spectrum and frequency, introduced in Appendix A can be used to construct standard tables of crossing rates.

4.1. The Rice Equation

A differentiable process \( x(t) \) is said to \textit{upcross} the level \( x_0 \) at time \( t_0 \) if there is a \( \delta > 0 \) such that the given realization of the process satisfies \( x(t_0 - \delta) < x_0 \) and \( x(t_0 + \delta) > x_0 \), with \( \dot{x}(t) > 0 \) for every \( t \in (t_0 - \delta, t_0 + \delta) \). \textit{Downcrossings} are defined similarly. A \textit{level crossing} can be either an upcrossing or a downcrossing of a given level, while the level crossing rate refers to the number of such crossings per time unit.

Rice \cite{89} determined the expected level crossing rate of a zero mean, continuous time, stationary, differentiable process \( x(t) \) in terms of the joint distribution of \( x(t) \) and its time derivative \( \dot{x}(t) \). That is, there should be

\[
\bar{N}(x_0) = \int_{-\infty}^{\infty} |\dot{x}| p_{x,\dot{x}}(x_0, \dot{x}) d\dot{x} = p_x(x_0) \int_{-\infty}^{\infty} |\dot{x}| p(\dot{x}|x_0) d\dot{x}
\]  

(4.2)

level crossings of the level \( x_0 \) on average per unit time. Here \( p_{x,\dot{x}}(x, \dot{x}) \) denotes the joint distribution of the process and its derivative, \( p_x(x) \) is the marginal probability distribution function (pdf) of \( x \), and \( p(x|\dot{x}) \) is the associated conditional distribution. Note that integrating from \((0, \infty)\) in \( \dot{x} \) results in the expected number of upcrossings. This formula assumes differentiability of the process, see e.g. \cite{30}, but fortunately this is not a problem for spectrum estimators which, as finite weighted sums of complex exponential functions, are entire functions of frequency.

Using this formula, Rice derived the expected number of zeros per second of a
zero-mean, stationary, continuous-time Gaussian process as

\[ N(0) = \frac{1}{\pi} \left[ -\frac{\gamma''(0)}{\gamma(0)} \right]^{1/2} = 2 \left[ \frac{\int f^2 S(f) df}{\int S(f) df} \right]^{1/2} \]  (4.3)

where \( \gamma(\tau) \) is the autocovariance function of the process at lag \( \tau \) and \( \gamma'' \) is the curvature of the autocovariance with respect to \( \tau \), and \( S(f) \) denotes the true spectrum. The bracketed term \( \gamma''(0) \) is also known as the variance of the derivative of the process (because, when they exist, the derivative and the expectation commute \[72\]), which arises because the conditional density in (4.2) is Gaussian with variance dependent upon that of the derivative of the process. For history of Rice’s formula \[88\] and an overview of the literature that has expanded from this result, consult \[14, 1\], or the excellent text by Cramér and Leadbetter \[30\].

### 4.1.2 Extension to \( \chi^2 \) Processes

Miller \[73\] and Silverman \[98\] independently extended the result of Rice to \( \chi^2 \) processes, see also \[53, 97, 9, 125, 66\]. The method of Barakat \[8, 9\], for the derivation of the joint density used in (4.2) is particularly accessible. Using the standardized spectrum estimate in (4.1), define

\[ \Upsilon(\Delta) = \text{Cov}\{z(f + \Delta), z(f)\}, \]  (4.4)

as the antecovariance of the spectrum estimate, where \( \Delta \) is the spacing in frequency and the antecovariance is \( \Upsilon(0) = 1/\alpha \). Denoting \( U(z; \alpha) \) as the upcrossing rate of the
level $z$ of a Gamma($\alpha$, $1/\alpha$) process per Rayleigh resolution\(^1\),

$$U(z; \alpha) = \psi \sqrt{\frac{z}{2\pi\alpha}} p(z; \alpha) \quad (4.5)$$

where $p(z; \alpha)$ is the Gamma($\alpha$, $1/\alpha$) pdf and

$$\psi = \frac{1}{N} \left[ -\frac{d^2}{d\Delta^2} \Upsilon(\Delta) \bigg|_{\Delta=0} \right]^{1/2} \Upsilon(0). \quad (4.6)$$

The factor of $\Upsilon(0)$ in the denominator converts from covariance to correlation, and the $1/N$ converts to frequency in units of Rayleigh resolution. The $-\frac{d^2}{d\Delta^2} \Upsilon(\Delta) \big|_{\Delta=0}$ term is the variance of the derivative $z'(f) = \frac{d}{df} z(f)$.

Under the null assumption that the data is white Gaussian noise, the upcrossing rate, e.g., $U(z_{99\%}; \alpha)$, where $0.99 = \int_0^{z_{99\%}} p(z; \alpha)dz$ gives the rate of false detection of periodic components in the spectrum which have over 99% significance. This false detection rate depends on the serial correlations $\Upsilon(\Delta)$ of the spectrum estimate, and these are derived for both classes of nonparametric spectrum estimators discussed in Appendix A, namely direct and multitaper estimators. Application of the above theory is applied with some care. It is usually assumed that the individual Gaussian processes making up the $\chi^2$ process are independent with identical autocovariances, but when applied to multitaper eigencoefficients neither of these assumptions is true.

\(^1\)Rayleigh resolution is defined as $R = 1/(N\delta t) = 1/T$ where $\delta t$ is the sampling rate and $T$ is the total duration of the series (in seconds, say, so that $R$ is in Hz). For a discussion of why one needs to standardize the units of frequency in this chapter as well as standardizing the spectrum, consult Appendix A.4.
4.1.3 Direct Spectrum Upcrossing Rate

The direct estimate of the spectrum is defined in Appendix A Eqn. (A.1) as the magnitude squared Fourier transform of zero-mean discrete data \( \{x_t\}_{t=0}^{N-1} \) multiplied by a data taper \( D_t \). Beginning with equations (4.5) & (4.6), and assuming a complex Gaussian noise process (the null hypothesis), the antecovariance, \( \Upsilon(\Delta) \), is (see also [110] or [114] §4.4),

\[
\Upsilon_D(\Delta) = \left| (D \ast D^*)(\Delta) \right|^2 = \left| \sum_{t=0}^{N-1} D_t^2 e^{-i2\pi t \Delta} \right|^2 \tag{4.7}
\]

where \( \ast \) used as a superscript denotes complex conjugate. The periodogram antecovariance is plotted in Fig. 4.3. To obtain the upcrossing rate, substitute \( \alpha = 1 \) so the distribution \( p(z; \alpha) \) of the standardized spectrum is \( \chi_2^2 \) or exponential, i.e. \( p(z; 1) = e^{-z} \). Then (4.6) becomes, upon application of Isserlis’s formula, for the fourth moments of a complex normal, Gaussian process

\[
\psi_D = \frac{2\pi}{N} \left[ \sum_{s,t=0}^{N-1} (s-t)^2 D_t^2 D_s^2 \right]^{1/2} \tag{4.8}
\]

and the upcrossing rate (4.5) is

\[
U_D(z; 1) = \frac{e^{-z}}{N} \left[ 2\pi z \sum_{s,t=0}^{N-1} (s-t)^2 D_t^2 D_s^2 \right]^{1/2} \tag{4.9}
\]
For example, straightforward substitution of the periodogram \( (D_t = 1/\sqrt{N}) \) estimate into the antecorrelation formula \((4.7)\) results in

\[
\Upsilon_p(\Delta) = \left| \frac{\sin N\pi\Delta}{N \sin \pi\Delta} \right|^2
\]

(4.10)

And the mean number of upcrossings of a level \( z \) for a periodogram per Rayleigh resolution becomes simply

\[
U_p(z) = \left[ \frac{\pi z}{3} \right]^{1/2} \left[ \frac{N^2 - 1}{N} \right]^{1/2} \approx \left[ \frac{\pi z}{3} \right]^{1/2} e^{-z}.
\]

(4.11)

Note that for complex Gaussian white data, \( \mathbb{E}\{z(f)\} = 1 \) for all \( f \) and all peaks are spurious, so the upcrossing rate should be as small as possible for high levels \( z \). The upcrossing rate, among direct estimates and for \( N \) fixed, only changes in the constant \( \psi_D \), which depends only on \( N \) and the data window \( D_t \). Ideally, then \( D_t \) is such that the upcrossing rate is as small as possible, but choosing the data taper so it minimizes \( U(z; \alpha) \) does not always yield a good spectrum estimate. As an extreme example, one can have \( \psi_D = 0 \) if one chooses the data taper as a Kronecker delta, but this taper results in a bad estimate, since the spectrum is calculated from one data point only. Among “good” choices for \( D_t \), see [50], one finds that the reduction in \( U^D(z; 1) \) as compared to \( U_p(z; 1) \) is marginal.

### 4.1.4 Multitaper Antecorrelation

The multitaper estimator of the spectrum \( \hat{S}(f) \) is defined in §A.2 as a weighted average of direct estimates of the spectrum \( S^{(k)}(f) \) computed with \( K \) Slepian tapers \( \nu_n^{(k)} \) and weighted by the frequency-dependent weights \( d_k(f) \) - see Eqn. (A.7), where
4.1. CROSSING PROBLEMS

$k = 0, \ldots, K - 1$, and $n = 0, \ldots, N - 1$. The antecovariance, $\Upsilon(\Delta)$, of a multitaper estimate is evaluated in [114]. Beginning with a white, Gaussian, complex-valued series, as before, and substituting the expected weights $d^2_k(f) = \lambda_k$ in the definition of the multitaper spectrum (A.7), the antecovariance becomes [8]

$$
\Upsilon(\Delta) = \text{Cov}\{z(f), z^*(f + \Delta)\} = \frac{1}{(2C_R)^2} \sum_{n,m=0}^{N-1} \left[ \sum_{k=0}^{K-1} \lambda_k v^{(k)}_n v^{(k)}_m \right]^2 e^{-i2\pi\Delta(n-m)} \tag{4.12}
$$

When $K \approx 2NW$, substituting Eqn. (A.2) into the bracketed term and using Mercer’s theorem as well as the orthogonality property of the Slepian sequences, we obtain the approximation

$$
\Upsilon(\Delta) \approx \frac{1}{(2C_R)^2} \sum_{n,m=0}^{N-1} \left[ \sin 2\pi W(n - m) \right]^2 \frac{2\pi}{\pi(n-m)} e^{-i2\pi\Delta(n-m)} \tag{4.13}
$$

$$
= \frac{1}{(2C_R)^2} \sum_{\tau=-(N-1)}^{N-1} (N - |\tau|) \left[ \sin 2\pi W\tau \right]^2 \frac{2\pi}{\pi\tau} e^{-i2\pi\Delta\tau} \tag{4.14}
$$

where the second equality uses the change of variables $\tau = n - m$. A few remarks are in order. (i) This expression is approximate because the bracketed term in (4.12) is a truncated spectral decomposition of (A.2). The approximation is accurate because eigenvalues become very small for $k > 2NW$. (ii) $\Upsilon(\Delta)$ drops rapidly to almost zero at $\pm 2W$, [114] Fig. 3. (iii) Real valued processes have half the degrees of freedom within $\pm 2W$ of the origin and the Nyquist frequency than they have at other frequencies. (iv) The standardizing factor $2NW$ is the trace of the sinc matrix in (A.2) and is the sum of the squared weights, $\lambda_k$ here. (v) Eqn. (4.14) is the Fourier transform of a product of a Fejér kernel and a triangular window (defined by $(N - |\tau|)$ for $|\tau| \leq N$ and zero otherwise).
Because the crossing rate \([4.5]\) depends on the curvature of the antecorrelation, we take the second derivative of \([4.14]\) at the origin, denoted \(\sigma_d^2\) here (which is the numerator of \(\psi\), \([4.6]\)),

\[
\sigma_d^2 = \frac{1}{N^2} \text{Var} \left\{ \frac{d}{df} \hat{z}(f) \right\} = -\frac{1}{N^2} \frac{d^2}{d\Delta^2} \Upsilon(\Delta) \bigg|_0
\]

(4.15)

\[
= \frac{1}{(NC_R)^2} \sum_{\tau=-N}^{N} (N - |\tau|) \sin^2 2\pi W \tau
\]

(4.16)

\[
= \frac{1}{(NC_R)^2} \sum_{n=0}^{N-1} \sum_{m=-n}^{n} \sin^2 2\pi W m
\]

(4.17)

\[
= \frac{1}{2C_R^2} \left\{ 1 - \left[ \frac{\sin 2\pi C_R}{N \sin 2\pi W} \right]^2 \right\}
\]

(4.18)

\[
\approx \frac{1}{2C_R^2} \approx \frac{2}{\alpha^2}
\]

(4.19)

and the scale factor \(1/N\) keeps frequency in units of Rayleigh resolution \((\delta t = 1\) here). Note that bringing derivative operators inside the covariance in \([4.15], [72\ Ch. 2]\), one sees that \(\sigma_d^2\) is the variance of the derivative of the spectrum estimate. The Fejér kernel is used to simplify \([4.17]\). The second term in \([4.18]\) is zero for nonzero positive integer and half-integer values of \(C_R\), and is otherwise approximately \([\sin(2\pi C_R)/(2\pi C_R)]^2\). The second approximation in \([4.19]\) holds exactly when \(\alpha = 2C_R\).

**Distribution of the Derivative of the MT Spectrum**

Incidentally, the distribution of the derivative of the standardized multitaper spectrum can be derived from the results just derived in the last section. From \([4.15]\), its variance is

\[
\frac{\text{Var} \{ \frac{d}{df} \hat{S}(f) \}}{N^2} = \frac{S^2(f) \text{Var} \{ \frac{d}{df} \hat{z}(f) \}}{N^2} = S^2(f) \sigma_d^2
\]

(4.20)
where the factor of $1/N^2$ converts the units to Rayleigh resolutions. If we write

$$\xi = \frac{z'}{z} = \frac{d}{df} \ln z$$

(4.21)

then one can write the joint distribution of $\xi$ and $z'$, and integrating out $\xi$ gives the marginal distribution of the derivative $z'$, as derived in [9],

$$p_{z'}(\xi; \alpha) = \frac{\alpha^2}{\sqrt{\pi} \Gamma(\alpha)} \left[ \frac{\alpha^2|\xi|}{2} \right]^{\alpha-1/2} K_{\alpha-1/2}(\alpha^2|\xi|)$$

(4.22)

where $K_{\nu}(\cdot)$ is Macdonald’s function (a modified Bessel function of the second kind), defined as

$$K_{\nu}(x) = \frac{\Gamma(\nu + 1/2)(2x)^\nu}{\sqrt{\pi}} \int_0^\infty \frac{\cos t}{(t^2 + x^2)^{\nu+1/2}} dt.$$  

(4.23)

For a direct spectrum estimator $\alpha = 1$ and the distribution above is the Laplace distribution.

### 4.1.5 Multitaper Spectrum Upcrossing Rate

Beginning with (4.5) and (4.6), direct substitution of the expressions (4.14) and (4.19), and letting $K = 2NW$ gives $\psi \approx (2/\alpha)^{1/2}$, one can conclude the main theorem of the chapter.

**Theorem.** A standardized multitaper spectrum (4.1) with $\alpha$ degrees of freedom, computed on a complex normal Gaussian process upcrosses the level $z$

$$U(z; \alpha) = \psi \sqrt{\frac{z}{2\pi \alpha}} p(z; \alpha); \quad \psi = \frac{2\pi}{N} \left[ \frac{\sum_{n,m=0}^{K-1} \lambda_j v_n^{(j)} u_m^{(j)} \lambda_k}{\sum_{k=0}^{K-1} \lambda_k^2} \right]^{1/2}$$

(4.24)
times per Rayleigh resolution, on average, where \( p(z; \alpha) \) denotes the Gamma(\( \alpha, 1/\alpha \)) probability density \( (A.10) \). By the arguments above, this is approximately equal to

\[
U(z; \alpha) \approx \sqrt{\frac{z}{\pi \alpha}} p(z; \alpha)
\]  (4.25)

where the approximation uses the truncated expression \( (4.19) \) in \( \psi \).

This formula is deceptively simple as it involves the approximation \( (4.14) \). Exact evaluation of \( (4.6) \) can be computationally inconvenient when \( N \) is large. Table \( B.1 \) gives values of \( \psi \) for various choices of \( C_R \), where \( N = 10^4 \) was used in the computation. Table \( B.2 \) gives a useful standardized table of \( U(z; \alpha) \) for common choices of \( C_R = NW, K = \alpha \), but see also the electronic supplement to [117].

It is common to average multitaper spectrum estimates from overlapped data blocks similar to the Welch method [127]. Assuming \( J \) non-overlapping segments, each with \( 2\alpha \) degrees of freedom, the crossing rate becomes

\[
U_J(z; \alpha) \approx \psi \sqrt{\frac{z}{2\pi \alpha J}} p(z; \alpha J) \approx \frac{1}{\alpha} \sqrt{\frac{z}{\pi J}} p(z; \alpha J)
\]  (4.26)

where the last expression uses the asymptotic value of \( \psi \).

**4.1.6 Simulation**

Combining the results of sections 4.1.3 and 4.1.5, Fig. 4.1(a) shows the periodogram estimate’s upcrossing rates against that of the multitaper, along with the result of a white Gaussian noise simulation where \( N = 10^5 \). Multitaper parameters were chosen as \( NW = 5, K = 8 \). Note that the multitaper curve is everywhere below that of the periodogram, indicating that at every level of significance, the multitaper estimator
4.1. CROSSING PROBLEMS

Figure 4.1: Panel (a) shows the multitaper $U(z_p; K)$ (blue, marked MT) versus the periodogram $U_p(z_p; 1)$ (green, marked P) estimates’ crossing rates per Rayleigh resolution with white complex-valued noise as input. Red dashed lines give upcrossing rates for simulated data ($N = 100,000$) for the respective estimates. The y-axis gives crossing rates for the level $z_p$ which is obtained by inverting the cumulative probability distribution at the level $100(1 - p)\%$, $p$ is the p-value for the test for a harmonic component at a given frequency and is plotted on the x-axis. For the multitaper, parameters $NW = 5, K = 8$ were chosen. Panel (b) shows the average dwell bandwidth of crossings for these estimators. These figures are reproduced from [117].

will give fewer significant spurious excursions. Note that the dominant term for the shape of $U(z; \alpha)$ is in the probability distribution $p(z; \alpha)$, which is exponential for a direct estimate (larger variance than multitaper). Thus the level $z$, on the top $x-$axis, is the inverse complimentary (1-cdf) cumulative distribution of the values on the bottom $x-$axis.

Fig. 4.1 partly explains the periodogram’s overly “grassy” appearance. This shows that the periodogram estimator produces a very large number of very narrow spurious peaks. This simple fact makes signal detection using periodograms very difficult indeed. Periodograms are normally both windowed and smoothed in order
to make up for the estimate’s high bias and variance, see [84], but this smoothing also widens and distorts genuine lines when present.

4.2 Shape and Width of Spurious Peaks

4.2.1 The Width of Spurious Peaks in Spectrum Estimates

The upcrossing rate formulae have a particularly simple relationship with the width of these peaks, see [86] equation (3.41). Define the dwell bandwidth, $w(z)$ as the average frequency difference from the point where the estimate crosses $z$ with a positive slope and the next frequency where the estimate returns to a level below $z$. For small $z$, this dwell distance can correspond to more than one peak, but for large $z$ the dwell band usually contains a single maximum, so peak-bandwidth is almost synonymous. The average number of such excursions times their mean bandwidth gives the total frequency range above $z$, the complimentary CDF (CCDF), $Q(z;\alpha)$, so $Q(z;\alpha) = U(z;\alpha)w(z)$.

For example, using (4.11) and noting that the CCDF is just $e^{-z}$, the average dwell bandwidth for spurious peaks in a periodogram is

$$w_P(z) \approx \sqrt{\frac{3}{\pi}z^{-1/2}}$$ (4.27)

Rayleighs. Similarly, combining (4.5) with the gamma CCDF gives the average dwell bandwidth for spurious peaks in the multitaper estimate, in Rayleighs,

$$w_{mt}(z) \approx \alpha \sqrt{\frac{\pi Q(z;\alpha)}{z P(z;\alpha)}}. \quad (4.28)$$

Where the ratio of the CCDF to the PDF in the above formula corresponds to the
4.2. SHAPE AND WIDTH OF SPURIOUS PEAKS

reciprocal hazard function. Fig. 4.1(b) gives a comparison of the average dwell bands for the two estimates, along with a white noise simulation as before. It is clear that the multitaper estimate gives relatively wider peaks than the periodogram, while the periodogram produces many very narrow ones. The bandwidth of the estimate, $NW = 5$ here, is also plotted for comparison.

4.2.2 The Shape of Spurious Peaks in Multitaper Estimates

Related to the dwell bandwidth problem is the question of the shape of random peaks in spectrum estimates. In the time-domain crossing literature, [30], continuous-time Gaussian processes are assumed to be well-approximated by a second-order Taylor series approximation during an excursion. Fig. 4.2(a) shows the average of the 50 largest peaks in a multitaper spectrum estimate of white noise, and while a quadratic approximation may be convenient, it does not appear from the figure to be correct.

Blachman [13] showed that for Gaussian processes, large excursions take the shape of the autocovariance function. Adapting this result to spectra is not strictly valid as spectra have $\chi^2$ distributions, but is asymptotically valid if one increases the degrees of freedom with $N$ (equivalently, fixes the bandwidth parameter $W$), and for spectra with $> 20$ df, the result applies reasonably well. Reconsidering (4.14), the antecorrelation $\Upsilon(\Delta)$ is dominated by a triangular window of width $2W$. Additionally, the antecorrelation is an entire function of frequency so its curvature is well-defined. The conclusion is that the shape of random peaks in multitaper spectra is triangular with width $\pm 2W$, and this is seen in Fig. 4.2(a) from the black dashed curve.

On the other hand, if the spectrum genuinely contains a $\delta$-function at a given frequency, the spectrum estimate takes on the shape of the spectral window [84, p.
Figure 4.2: In panel (a), the red curve gives the average shape of the 50 largest (spurious) peaks in a multitaper white-noise spectrum. The light grey band shows the range, while the dark grey band shows one standard deviation away from the mean. Here, \( N = 100,000 \), \( NW = 5 \), and \( K = 8 \) was used. Dashed vertical lines give the \( \pm W \) and \( \pm 2W \) bands, and the horizontal dashed lines give the mean, \( 1 \), \( \pm 1\sigma = \pm \sqrt{1/K} \) and the 99% significance level. The dashed black line is the scaled theoretical antecorrelation for the multitaper estimator. At 99% significance, the average peak width is about \( W \) Rayleighs wide. In panel (b), a process containing 55 sinusoids in noise, spaced equidistant in frequency is simulated and the average shape of these peaks in an adaptively weighted multitaper spectrum is shown. The dashed black line is the magnitude squared spectral window where equal weighting was used. It is clear that a peak in a multitaper spectrum takes on the shape of the spectral window. These figures are reproduced from [117].

207], seen in Fig. 4.2(b)\(^2\) which, for a multitaper estimate is roughly rectangular with width \( \pm W \) centered about the frequency of interest. To see this, recall from appendix A that for a direct spectrum estimate (A.1), \( E\{\hat{S}^D(f)\} = |D(f)|^2 \ast \delta(f) \), where \( D(f) \) is the Fourier transform of the data window \( D_t \), and \( S(f) \) is a delta function when \( x(t) \)

\(^2\)The spectral window is the squared magnitude of the Fourier transformed data taper used to compute the spectrum. In the case of the periodogram, it is the Fejér kernel, [84]. For the multitaper, is the squared magnitude of a weighted average of Slepian functions, which are Fourier transformed Slepian sequences.
Figure 4.3: Using a similar color scheme to Fig. 4.2, the red curve gives the average shape of the 50 largest (spurious) peaks in a periodogram estimate of Gaussian white-noise. Similar to the last example, the shape of a true peak in a periodogram is expected to match the shape of its spectral window, the Fejér kernel (dashed black line), see e.g. [84, p. 198-200].

is a perfect sinusoid. Since the multitaper spectrum is a weighted average of direct estimates, $\hat{S}^{(k)}(f)$, \([A.6]\), the expected value near a genuine sinusoid is a weighted average of Fourier-transformed dpss tapers, i.e., the spectral window.

To illustrate this, the same experiment was performed using the same values of the parameters, this time adding 55 sinusoids separated equally in frequency $18W$ away from zero and the Nyquist frequency, with unit amplitude to the white noise from before, where the noise variance is adjusted to $(1/16)$. The result of this experiment is shown in Fig. 4.2(b). Note that here squared adaptive weights have been chosen simply as $1/K$ for all $f$, so the black dotted line in panel (b) is the spectral window using equal weights.

The shape of a spurious peak in a periodogram estimate likewise behaves like its
4.3. Example

Theoretical antecorrelation, see Fig. 4.3. The antecorrelation $\Upsilon^P(\Delta)$ (black dashed curve) drops to near zero past $1R$, which demonstrates that periodogram estimators are nearly uncorrelated when their frequencies are separated by more than one Rayleigh resolution. Returning to the width of excursions shown in Fig. 4.1, one gets an interesting comparison. Note that different choices of the multitaper bandwidth parameter, $W$, result in wider or narrower peaks.

4.3 Examples

In this section I will revisit the data presented in the last chapter to show how the results developed in this chapter are useful. To use the results developed in the last few sections on real data, I use a permutation test. Upcrossings in spectra computed from these data as well as spectra computed using a temporally permuted version of the data are counted and compared. Temporal permutation of the data sequence destroys the original covariance structure in the original data, and hence may be expected to have a number of upcrossings similar to that expected for Gaussian white noise. The comparison between the expected number of peaks in these spectra and the number they actually have is of interest, as the time-domain versions of these data are frequently highly non-Gaussian. Refer to the tables in Appendix B for expected numbers of peaks in Gaussian noise.

4.3.1 ACE Proton Density

Recall the solar wind data presented in section 3.2. There it was shown that the spectrum in the $2 - 3$ mHz band could be decomposed into approximately half modal and non-modal components and the raw data had non-Gaussian moments. Now the
4.3. EXAMPLE

Table 4.1: Crossing table for the ACE data 2000 – 3000µHz, see text. The column marked $P$ represents the significance level, $z$ is the inverse CDF of the standardized spectrum’s distribution under $H_0$ at the significance level $P$. $U(z; \alpha)$ corresponds to Eqn. (4.5) times $10^6$, i.e., the number of crossings of the spectrum per $10^5$ Rayleigh resolutions, while $\sqrt{U(z; \alpha)}$ is the expected standard deviation for a counting process. The column marked Data is the number of crossings observed on $10^5$ Rayleigh resolutions in the band $2000 – 3000\mu$Hz of the ACE standardized spectrum, Fig 3.3, and Permuted refers to the number of crossings one obtains when one first permutes the temporal order of the data. Note the extreme number of peaks - 61 peaks above the 99.999% point of the spectrum when one only expects 1.

<table>
<thead>
<tr>
<th>$P$</th>
<th>$z$</th>
<th>$U(z; 6)$</th>
<th>$\sqrt{U}$</th>
<th>Data</th>
<th>Permuted</th>
</tr>
</thead>
<tbody>
<tr>
<td>50</td>
<td>0.945</td>
<td>14584</td>
<td>121</td>
<td>12453</td>
<td>15536</td>
</tr>
<tr>
<td>90</td>
<td>1.546</td>
<td>5940</td>
<td>77</td>
<td>7645</td>
<td>6253</td>
</tr>
<tr>
<td>99</td>
<td>2.185</td>
<td>861</td>
<td>29</td>
<td>2399</td>
<td>943</td>
</tr>
<tr>
<td>99.9</td>
<td>2.742</td>
<td>106</td>
<td>10</td>
<td>697</td>
<td>111</td>
</tr>
<tr>
<td>99.99</td>
<td>3.261</td>
<td>12</td>
<td>3</td>
<td>201</td>
<td>12</td>
</tr>
<tr>
<td>99.999</td>
<td>3.756</td>
<td>1</td>
<td>1</td>
<td>61</td>
<td>1</td>
</tr>
<tr>
<td>99.9999</td>
<td>4.235</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
</tr>
</tbody>
</table>

question is whether the perceived modal structure in this spectrum is genuinely due to the covariance structure of the process or the non-Gaussian distribution of the data, which can be tested using a permutation test.

Table 4.1 shows the number of crossings of the standardized spectrum, normalized per $10^5$ Rayleigh resolutions, of the 50% – 99.9999% percent level of the central distribution. For this spectrum, the Rayleigh resolution is 2.15nHz, and there are thus 464,896 Rayleighs on the 2000 – 3000µHz band. Table 4.1 shows a pathological number of peaks in the original data (column marked data), many more than expected (column marked $U(z; 6)$). The standard deviation given by assuming Poisson statistics for the crossings is in the column marked $\sqrt{U(z; 6)}$, (asymptotic result). Comparison with the expected number and standard deviation, it is clear that there are many more peaks than one would expect were the process free of line components.
Given that the data are non-Gaussian in the first place, one might question whether the distribution of the data itself could be producing spectra with large numbers of peaks. One might also wonder whether the variance on the crossing rate expression could be large. In order to rule out these possibilities, the original data are temporally permuted, spectra computed, and then the number of crossings observed per \(10^5\) Rayleighs is given in the last column of Table 4.1. The permuted data is seen to have approximately the same number of peaks as expected for a Gaussian white noise process. The conclusion to be made from this experiment is that it is specifically the autocorrelation structure of the process, equivalently the true spectrum, which is producing the appearance of a myriad of peaks.

### 4.3.2 Example: Neutron Data

<table>
<thead>
<tr>
<th>(P), (%)</th>
<th>(z)</th>
<th>(U(z; 6))</th>
<th>(\sigma)</th>
<th>Thule Data</th>
<th>Thule RP</th>
<th>Newark Data</th>
<th>Newark RP</th>
<th>South Pole Data</th>
<th>South Pole RP</th>
</tr>
</thead>
<tbody>
<tr>
<td>50.000</td>
<td>0.959</td>
<td>13097</td>
<td>114</td>
<td>8413</td>
<td>13075</td>
<td>8548</td>
<td>12856</td>
<td>12141</td>
<td>12987</td>
</tr>
<tr>
<td>90.000</td>
<td>1.471</td>
<td>5386</td>
<td>73</td>
<td>5511</td>
<td>5109</td>
<td>5093</td>
<td>5202</td>
<td>7229</td>
<td>4711</td>
</tr>
<tr>
<td>99.000</td>
<td>2.000</td>
<td>784</td>
<td>28</td>
<td>1610</td>
<td>680</td>
<td>1658</td>
<td>745</td>
<td>1740</td>
<td>646</td>
</tr>
<tr>
<td>99.900</td>
<td>2.453</td>
<td>97</td>
<td>10</td>
<td>446</td>
<td>116</td>
<td>328</td>
<td>86</td>
<td>390</td>
<td>102</td>
</tr>
<tr>
<td>99.990</td>
<td>2.870</td>
<td>11</td>
<td>3</td>
<td>124</td>
<td>11</td>
<td>97</td>
<td>11</td>
<td>73</td>
<td>11</td>
</tr>
<tr>
<td>99.999</td>
<td>3.265</td>
<td>1</td>
<td>1</td>
<td>29</td>
<td>2</td>
<td>27</td>
<td>2</td>
<td>9</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 4.2: Results of a permutation test between 2-3mHz on the three pressure corrected neutron data sets. The column marked \(P\) denotes the significance level, \(z\) the level of the standardized spectrum for that significance, \(U(z; 6)\) is the expected number of upcrossings per \(10^5\) Rayleigh resolutions for Gaussian white noise, and \(\sigma\) is the standard deviation assuming Poisson statistics for the upcrossings. The columns marked Data and RP denote the number of upcrossings of the data and of the randomly permuted version of the data. As one can see, randomly permuted data possess numbers of crossings similar to what one would expect for Gaussian noise.

To test whether the number of peaks seen in the neutron data are a result of the
distribution of the data, the multitaper estimator or the data’s correlation structure, one can use a permutation test on the middle of the $p$–mode band $2−3$mHz. In table 4.2 the column $P$ denotes the significance level, $U(z)$, $\sigma$ denote the expected number of upcrossings of that level (or peaks) per $10^5$ Rayleigh resolutions of a multitaper spectrum using $NW = 3.5$ and $K = 6$, and the standard deviation expected with Poisson crossing statistics. At $\geq 99.99\%$ significance, the number of peaks observed can be more than $20\sigma$ times expected. This supports the hypothesis that cosmic radiation reaching Earth contains modal signatures.

### 4.3.3 Example: Surgical Scheduling

In a clinical anesthesiology paper by Moore et al. [76] 6-years of surgical demand time series were decomposed into periodic, autoregressive, and linear components. The goal of the analysis was specifically signal detection and estimation for the purpose of reconstruction and eventual prediction of surgical demand. The main result of the study was that 54 periodic components (detected not using a multitaper spectrum, but a related tool, the harmonic F-test, §VIII of [111]) significant at 99% were sufficient to explain 76% of the variance in the series. The main residual after the removal of these linear and periodic terms was recognizable as holiday variance, see Fig. 7 of Moore et al. [76].

Table 1 of the study of Moore et al. shows that for this series 83, 54 and 30 detections were observed in a harmonic F-test (see Appendix A.5.1) above the 98, 99, and 99.5% significance levels, respectively. The authors expect that in time and frequency domains, the data have the same degrees of freedom, so for 2,191 daily samples, the number of significant detections above 98, 99, and 99.5% should be
approximately 44, 22, and 11, respectively. The discrepancy here is large - more than double the peaks expected are observed. But these observations spur more investigation - how many peaks ought one to observe, on average, given an F-test of data containing no line components?

4.4 Extensions to the crossing theory

Given that this chapter has suggested a new approach to spectrum analysis of time series which appear to contain “many lines”, what are the next steps? The easiest extension may be to the harmonic $F$-test for periodic components. In this section I will present an asymptotic result and some simulations showing that the crossing rate of the F-test is, up to some scaling factor, the same as for a $\chi^2$ distributed spectrum estimate, such as the periodogram.

4.4.1 Harmonic F-test

The harmonic F-test is a signal detection method used with multitaper estimates, introduced in Appendix A, and in this section I will discuss the crossing rate of significant excursions of such an F-test.

Number of Crossings

The crossing rate of the F-test statistic is surprisingly difficult to derive. The distribution of the derivative of the F-statistic is easily seen to be zero-mean normal, when conditioned on the value of $F$ and its denominator $D$. The variance of the derivative of the standardized eigencoefficients however, are different for every pair $j,k$. This

\footnote{This expected value is rough, see the comments in section 5 of \textit{[117]}.}
makes manipulation of the variance expression difficult to put in terms of identifiable quantities. If one does assume that these quantities are equal for different pairs \( j, k \), say \( \dot{\sigma} \), (which holds under certain limiting assumptions), then the variance of the \( F \)-test is

\[
\text{Var}\{\hat{F}|F,D\} = \frac{4F\dot{\sigma}^2}{D} [(K - 1) + F].
\] (4.29)

and using a derivation similar to that of Barakat [8], but conditioning on both \( F \) and \( D \) and eventually integrating to remove the dependency on \( D \), one arrives at the following formula

\[
U_F(F; K) \approx \frac{\dot{\sigma} \sqrt{2F(2(K - 1) + 2F)}}{2\sqrt{\pi}(1 + 2F/(K - 1))^{K-1/2}} \cdot \frac{\Gamma(K - 1/2)}{\Gamma(K)}
\] (4.30)

Letting \( N \) grow large while \( W \) is fixed, that is, as \( K \) increases, the distribution of \( 2F \) approaches that of a \( \chi^2_2 \) random variable, and substituting equation (4.29) for \( \dot{\sigma}_z \), which is approximately \( 2\sqrt{\pi} \), in the limit gives

\[
U_p(z) \approx \lim_{K \to \infty} U(F; K) = \frac{\dot{\sigma}_z}{2\sqrt{\pi}} \sqrt{2} e^{-z} = \sqrt{2} e^{-z}.
\] (4.31)

Fig. [4.4] shows the behavior of the crossing rate of the F-test statistic as \( K = 2NW \) grows. The solid black line shows the asymptotic crossing rate, while the dashed lines, in RGBIV order denote \( N \) increasing. To conclude this section, the crossing rate of the F-test statistic with 2 and \( 2K - 2 \) degrees of freedom behaves, in the limit, like the periodogram.
Figure 4.4: Simulated white noise process peaks in the $F$-test. Here $N = 1,000$ (red), 5,000 (green), 10,000 (blue), 100,000 (purple), $K = NW$ and $W = 1/1000$ was used so that $NW = 1, 5, 10, 100$. As $N$ grows large, it appears that the number of peaks approximately approaches that of a standardized periodogram estimate - solid black line.

4.5 Solar Modes and Multiple Comparisons

This chapter has answered questions about the number of peaks expected in spectra and the distribution of the spectrum under the null hypothesis. However, the “elephant in the room” remains the issue of multiple comparisons. How does one identify individual modes in a spectrum like that of the ACE data, in which there are 10 million modes expected in the $p$-mode band [51] p. 33], and rotational splitting divides each mode into $2l + 1$ equally spaced singlets? Does one expect to detect each singlet with high significance? What if some singlets are missing? Can one use something other than the rate of false detection [10] [11]?

A related question is “how many ways can one misidentify a mode?” For example,
the asymptotic formula (2.3) shows that it is easy to misidentify a \( p(l, m, n) \) mode with center frequency \( \nu_{l,m,n} \) for a \( p(l + 2, m, n - 1) \) mode at approximately the same center frequency. The \( p(l, m, n) \) mode will have \( 2l + 1 \) singlets equally spaced by about 440nHz and the \( p(l + 2, m, n - 1) \) mode will have \( 2l + 5 \) singlets. If only, say, 3 singlets are observed, how can one be sure that one has found an \( l = 1 \) mode vs an \( l = 2 \) mode with four non-significant outer singlets?

Conventional multiple comparisons approaches require that the number of possible hypotheses to be tested is known in advance, however this does not seem to be true for spectrum estimates. With spectrum estimates, contiguous frequency bins are correlated, and the expression for the approximate number of peak positions depends upon the expected dwell band. Suppose that a bandwidth of \( B = 1000R \) is given. Then the expected number of spurious peaks is \( p = B \cdot U(z; \alpha) \), which can be distributed over about \( q = B/D(z) \) positions. The number of ways one can rearrange \( p \) spurious peaks in \( q \) positions is \( \binom{q}{p} \). If one uses a multitaper estimate with \( C_R = 5 \) and \( K = 8 \) and selects the 90% significance level, \( p = 54 \), \( q = 1000/1.857 \), and \( \binom{q}{p} = 8.9 \times 10^{74} \) possible combinations. This staggering number of combinations makes correction for multiple comparisons almost impossible. One can see from [117, Fig 3(a) and Table 2] that the 75% level is close to the maximum of the number of false patterns.

The point is that the event of detecting a mode with equally spaced singlets is an extremely rare occurrence, combinatorially speaking, and becomes rarer the larger the number of singlets that appear and the larger the chosen significance threshold. The expressions derived in this chapter do not give a simple formula for correction for the massive number of tests for lines one does when one computes a standardized
4.6 Conclusion

Expressions for the upcrossing rate of the spectrum of a Gaussian white noise process above a given significance level have been given in this chapter. This expression involves the autocorrelation of the spectrum, or antecorrelation, as well as the estimator’s probability distribution. Among nonparametric spectrum estimates, fewer spurious detects are always given by the multitaper estimate. From the upcrossing expression one also can derive the expected width of excursions above a given significance level. The shape of the spectrum near a spurious peak corresponds to the shape of the antecorrelation there. Because the multitaper antecorrelation is roughly triangular in shape with width $\pm 2W$, spurious peaks take on that shape, while genuine peaks take the shape of the spectral window, which is rectangular with width $\pm W$.

The question of spectra containing spurious line components is not a new one, and even in 1898, Schuster considered *The Separation of Real from Accidental Periodicities*, [96], an important consideration in spectrum estimation. Estimation of the significance of perceived line components in spectrum estimates was considered by Fisher [39], and until now, these questions have stymied those working with time series. Especially when time series appear to contain excessive numbers of lines, [60], the validity of spectrum estimation itself is often questioned, instead of entertaining the real possibility of that many lines in noise. The results of this chapter attempt to answer these questions.

A permutation test was used to justify the perceived modes in the ACE SWEPAM
data and in the neutron data. After random permutation, the large number of up-crossings of the spectrum at high significance levels reduces to the number expected for a Gaussian process. This evidence, combined with the previous chapter’s results that these spectra are mainly modal, supports the argument that these peaks are genuine. That is, these data contain peaks which are indeed due to the correlation structure of the data and not its distribution. The agreement of the modal structures with known modal frequencies in the $p$–mode band supports these claims.

The particular application to the neutron monitor data is entirely novel, and suggests solar modulation of charged particles traveling at relativistic speeds can have a noticeable effect on time scales much shorter than the 11-year solar cycle. This conclusion may prove to be of considerable scientific interest.

In conclusion, the present approach is part of an effort to better characterize spectra of this sort. Processes with “many lines” are certainly not scarce in nature, and diverse examples exist. In particular, the example with the solar modes is a prominent one; it is likely that these modes manifest themselves in many other natural processes in space and on Earth.
Chapter 5

Parametric Modeling of Solar Oscillation Spectra

*Peak bagging* is a term used to describe the identification of complicated spectral line profiles in noise, similar to signal detection, but with application in this case to data from images of the solar disk. Looking ahead to the with model equations (5.1) for the oscillations, the likelihood function near each solar oscillation frequency depends upon the center frequency, \( \omega_0 \), the width, \( \alpha \), and the amplitude, \( A \), of the mode as well as the noise background \( \sigma^2 \).

For completeness, this chapter begins in §5.1 with a literature review of the models used to describe the spectral characteristics of optical Doppler data, as described in §3.1 on GONG disk images, and is a continuation of the discussion begun in §2.5.

I will then describe a simple maximum likelihood method, borrowed from the communications literature, for estimating the parameters of solar modes. The reason for the simplicity of the model is that the intended application is not helioseismic data, where much is known about the spectra of solar disk image time series, where individual modes can be isolated, but composite spectra such as the ACE data or the neutron monitor data.
5.1 Literature Review: Modeling of Solar Oscillations

Global $p$–mode frequencies have been determined from optical Doppler measurements by various authors. Beginning chronologically, there are a number of papers, [4, 94, 6], which describe a maximum likelihood-based parameter estimation process for these spectra. I’ll begin with the simplest of these, the model of Anderson, Duvall and Jefferies.

5.1.1 The Model of Anderson, Duvall and Jefferies

The simple model of Anderson, Duvall and Jefferies [4] uses the following second order ordinary differential equation for harmonic oscillators, assuming the modes are stochastically excited and intrinsically damped, using angular frequency $\omega = 2\pi f$ for convenience,

$$\frac{1}{\omega_0^2} \frac{d^2}{dt^2} x(t) + \frac{1}{\omega_0 Q} \frac{d}{dt} x(t) + x(t) = y(t) \tag{5.1}$$

where $x(t)$ is the displacement from equilibrium, $\omega_0$ is the angular frequency of the undamped oscillator, and $y(t)$ is the forcing function, which is usually assumed to be real white noise. The quality factor, $Q$, describes the damping of the system: when $Q > 1/2$ the system is underdamped, $Q = 1/2$ is critically damped and $Q < 1/2$ is overdamped. The $Q$ of a mode represents $2\pi$ times the ratio of the total energy in the mode divided by the energy lost per cycle. $G$–modes are known to stay relatively stable over long periods and thus have high $Q$. The width at half-power, $\alpha$, of the oscillator is related to the quality factor by $\alpha = \omega_0/Q$, which for an underdamped system represents the rate of exponential decay, or damping, of the oscillations.

This stochastic differential equation is solved by taking the Fourier transform of
Figure 5.1: Panel (a) shows transfer function profiles, see Eqn. 5.3. For the solid curve, only one peak is present with $f_0 = 0.2$, $Q = 10^3$, $\sigma^2 = 0$ and $A = 1$. The dashed curve is a superposition of peaks at $f_0 = 0.1, 0.4$ and $Q = 10^3, 10^3$ where the first peak has four times the amplitude. The dotted curve has $f_0 = 0.25, 0.05, 0.3$ and $Q = 10^3, 10^3, 10^2$ with amplitude 1. For each sequence, $N = 500$ frequency bins were used. In part (b), the solid curve from (a) is shown again along with its scaled Lorentzian approximation (dashed curve). Note that though the Lorentzian is symmetric, it reproduces the transfer function very well near the peak.

both sides, recalling that $-i\omega X(\omega)$ is the Fourier transform of $x'(t)$ when $X(\omega)$ is the Fourier transform of $x(t)$

\[ X(\omega) = H(\omega)Y(\omega) \]  

(5.2)

letting $Y(\omega)$ and $X(\omega)$ denote the Fourier transforms of $y(t)$ and $x(t)$ respectively, and $H(\omega)$ denotes the transfer function of the system, which is

\[ H(\omega) = \omega_0^2 \frac{\omega_0^2 - \omega^2 + i\alpha\omega}{(\omega_0^2 - \omega^2)^2 + (\alpha\omega)^2}. \]  

(5.3)

Fig. 5.1 shows the overall shape of the magnitude squared transfer function $|H(\omega)|^2$ for various $\omega_0$ and $Q$. The power spectrum of $x(t)$ near a resonance $\omega \approx \omega_0$, is just
\( |H(\omega)|^2 |Y(\omega)|^2 \), which, assuming Gaussian noise for \( y(t) \), and since \( p \)-modes tend to have high \( Q \) and narrow width,

\[
L(\omega) = \frac{A\alpha}{(\omega - \omega_0)^2 + (\alpha/2)^2 + \sigma^2}
\]  

(5.4)

where \( A \) is the integrated power of the mode and \( \sigma^2 \) is the noise variance. Equation (5.4) is the Lorentzian profile for the mode.

When using disk-integrated Doppler velocity observations, the signal exhibits a superposition of modes. Anderson et al. began with the Lorentzian and the exponential 1 distribution of standardized direct spectrum estimate at each frequency and defined a parametric model for the mean value of the spectrum.\(^1\) Because periodogram estimates are approximately uncorrelated when separated by a Rayleigh or more in frequency, see Fig. 4.3, the likelihood in a band was the product of exponential-1 densities at frequencies spaced one Rayleigh apart. The model \( M_i(\mathbf{a}) \), where \( i \) indexes the equally-spaced frequency bins, is used to standardize the periodogram (that is, it is used as the ideal mean value function of the spectrum as a function of frequency), where the model resembles (5.4),

\[
M_i(\mathbf{a}) = \sum_{j=1}^{N} \frac{A_j(\alpha_j)^2}{(\omega_i/\omega_{0j})^2 + (\alpha_j)^2} + \sum_{k=1}^{n} c_k \omega_i^{k-1}
\]  

(5.5)

and the parameters \( \mathbf{a} \) are: \( A_j \), the maximum signal power, \( \alpha_j \) is the half width at half-maximum, \( \omega_{0j} \) is the angular mode frequency, \( N \) is the number of modes in the

\(^1\)Although these authors used the periodogram spectrum estimator to derive their results, the introduction of \( 4 \) laments the inconsistency of the periodogram: “The measurement of the line-profile parameters proves to be difficult for several reasons: (1) the erratic nature of the spectrum, which does not decrease with an increase in the length of the observation, is given to concealing the characteristics of the line profiles;” see also Fig. 1 of \( 4 \).
fitting interval, and $c_k$ are the coefficients describing the background power, see Eqn (15) in [3]. Assuming independence of the individual modes (because direct estimates are uncorrelated when separated by a Rayleigh or more in frequency) the likelihood on the band of interest is just a product of $\chi^2$ (exponential) densities,

$$L(\{\hat{S}(\omega_i)\}_i) = \prod_i M_i(a)^{-1} \exp \left\{ -\frac{\hat{S}(\omega_i)}{M_i(a)} \right\}$$  \hspace{1cm} (5.6)

When the logarithms of the amplitude $A$ and width parameters $\alpha$ are used, the likelihood function for the mode parameters becomes approximately Gaussian in distribution so that the inverse of the Hessian matrix, as in the usual multiple regression problem, can be used to produce estimates of the uncertainty for the parameter estimates.

Anderson et al. used a downhill simplex algorithm to maximize the likelihood and determine the associated parameters. In fitting mode frequencies, adequate resolution is key, and the authors remark that when spectra with inadequate resolution are used, the simplex method will select large narrow peaks while also producing unreliable error estimates.

5.1.2 M-averaged spectra

If one assumes that one can treat the spectra for each $l$ as independent realizations of the same process near each $m$—value (see §3.1), then one can obtain a power spectrum with more degrees of freedom (and hence “better behaved” statistics) by averaging.

---

²The bandwidth of the periodogram is (without accounting for spectral leakage) twice the width of the main lobe of the Fejér kernel [84]. It is interesting that, given the results in Chapter 4 on the number of false large peaks in periodogram estimates, modal parameters were able to be reliably estimated at all from these spectra.
5.1. REVIEW: MODELING SOLAR OSCILLATIONS

the spectra obtained near each $m$. That is, averaging the spectrum obtained at $2l+1$ places

$$
\hat{S}_{m-\text{ave}}(f; \hat{f}_0) = \frac{1}{2l+1} \sum_{m=-l}^{l} \hat{S}_x(f + a_1 \cdot m - \hat{f}_0)
$$

(5.7)

where $a_1$ is the optical splitting observed (usually $\sim 440nHz$ for $p$-modes), and $\hat{f}_0$ is within 40nHz of the center frequency determined optically, $f_{\text{opt}} \pm 40nHz$. The resulting $m$-averaged spectrum $\hat{S}_{m-\text{ave}}(f; \hat{f}_0)$ has a $\chi^2_{(2l+1)\nu}$ distribution, where $\nu$ is the number of degrees of freedom of the spectrum estimate. The use of this technique is obviously not appropriate when the bandwidth of the spectrum estimate is close to $a_1$. One also forgoes any knowledge about the shape of the individual $m$-spectra by averaging, but it can allow for better detection when the signal to noise ratio is low.

5.1.3 Spectral Line Asymmetry and the Nigam-Kosovichev model

More recently the model (5.1) has been adapted to include a slight line asymmetry observed in helioseismic power spectra, described by e.g. Nigam and Kosovichev [78], which is caused by interferences between excited and reflected waves. It was recognized that a symmetric model was not appropriate for the solar oscillations because the two components of the spectrum, due to the solar oscillation(s) (signal) and granulation (noise) were significantly correlated. The result of fitting an asymmetric line with a symmetric model results in significant shifts of the center frequency. The asymmetric model was able to explain differences in sign between VIRGO and MDI asymmetries [121] and thus had considerable predictive power. The model equation includes an asymmetry parameter $B$ and the power spectrum is given by an asymmetric Lorentzian

$$
S(g) \approx A \frac{(1 + Bg)^2 + B^2}{1 + g^2} + B_l
$$

(5.8)
where \( g = 2(\omega - \omega_0)/\alpha \), and \( B_l \) is the uncorrelated linear background. The parameters to be fit are \( A, B, \alpha \) and \( f_0 \). The asymmetry parameter \( B \) is positive for positive asymmetry (more power to the high frequency end of the peak) and negative for negative asymmetry. When the noise is uncorrelated with the oscillation, \( B = 0 \).

Nigam and Kosovichev used a minimum mean squared error criterion in order to fit the mode parameters to observational data from SOHO (MDI) as well as simulated data.

**The Complex Gaussian Model of Schou**

A different approach to the likelihood is given in the PhD dissertation of Schou \[94, appendix\], which is summarized and extended in the papers of Appourchaux *et al.* \[6\], which bases the likelihood on the (complex) Fourier-transformed data instead of the estimated spectrum. This approach is useful because phase information, ignored in all the models up to this point, is incorporated and correlations between solar oscillations for different \((l, n)\) are not ignored. Additionally, Fourier-transformed data, together with the fact that solar disk images contain large numbers of convective granules, is reliably complex Gaussian distributed \[68, 19\] with uncorrelated real and imaginary parts. That is, for a particular mode with frequency \( f_i \) the density is

\[
p(a, f_i) = \frac{1}{\sqrt{2\pi v(a, f_i)}} e^{-\frac{1}{2}y(f_i)^2/v(a, f_i)}
\]  

(5.9)

for both the real and imaginary parts, where \( a \) is a vector of parameters to be estimated, \( y(f_i) \) are the (tapered) Fourier-transformed data at frequency \( f_i \), and \( v(a, f_i) \) describes the covariance structure. The expression for the likelihood, however, must be designed to incorporate the strong correlations between individual modes, which
is the result of the fact that the expansion in spherical harmonics, though orthogonal over the entire surface of the Sun, is not orthogonal over the surface visible at any given time. The modes are not truly isolated by the spatial mask, so the Fourier transform for \((l, m)\) may be a sum over a number of different modes. Letting

\[
z_n(f_i) = \sum_k c_{n,k} y_k(f_i)
\]

(5.10)
denote the Fourier transform of the observed time series, where the contributions of the individual mode Fourier transforms \(y_k\) are given weights \(c_{n,k}\). The matrix with entries \(c_{n,k}\) is called the leakage matrix \([95]\) and is estimated from the known quantities relating the modes, as well as atmospheric and instrumental conditions. Letting \(K\) denote the covariance matrix of the \(z_n\)'s, i.e.

\[
K_{n,m}(a, f_i) = \text{Cov}\{z_n(f_i), z_m(f_i)\} = \sum_k c_{n,k} c_{m,k} v_k(a, f_i)
\]

(5.11)
so the probability density at frequency \(\nu_i\) is

\[
p_{z}(a, f_i) = \frac{1}{\sqrt{2\pi |K(a, f_i)|}} e^{-\frac{1}{2}z_i^T K(a, f_i)^{-1}z_i}
\]

(5.12)
one can minimize the associated negative log likelihood function either by differentiating directly, or producing contour plots of the likelihood associated with each choice of the parameters \(a\) \([94]\) Eqns (E10)-(E20) appendix].
5.1.4 The state-of-the-art for solar mode parameter fitting

The method of Schou, applied to the asymmetric model equation of Nigam and Kosovichev, with or without a maximum likelihood-type optimality criterion or \( m \)-averaging forms the basis for modern mode-fitting methodology. Some additional variations and innovations have been presented by the following authors:

1. Korzennik [62] uses a multiple sine taper spectral estimator [90] together with asymmetric Lorentzian profile and a least-squares optimization for the parameter values. Korzennik provides, on his personal website, a list of fitted mode frequencies from together with their splittings from long time series from both GONG network and the next-generation MDI (HMI) data.

2. Salabert et al. [92, 91] is able to measure low signal to noise ratio \( p \)-modes as a result of \( m \)-averaging over individual spectra using the asymmetric model. The latter paper, which gives \( p \)-mode frequencies in the \( 1 - 2 \)mHz frequency range obtained from GONG and MDI data is of particular interest, because the \( p \)-modes in this range are particularly difficult to detect because of low signal to noise ratio in this range.

3. A “Pseudo-global” method of mode fitting was put forth by Fletcher et al. [41]. This methodology was used by Broomhall et al. on data from the BiSON network to generate modal frequencies for the varying levels of solar activity (the BiSON network has been gathering resonance scattering spectrometer data for over 30 years) [20].

\(^3\)Korzennik’s website is located at \texttt{https://www.cfa.harvard.edu/~sylvain/}.
The development of methods for parameter fitting of Doppler-velocity images has been ongoing for over 25 years, and the current state of the art in mode fitting can generate fits to these oscillations with little uncertainty.

However, the task at hand is to deal with composite spectra, that is, data which appear to contain all of the modes. In Chapter 3 I dealt with the problem of composite spectra by characterizing the empirical distribution of the spectrum estimate by fitting a mixture model, and in Chapter 4, I determined the false detection rate for periodic components in noise. The goal of this chapter, then, is to begin with a simple model, combine it with complex demodulation, to exclude frequency components which lie far-away from the oscillations of interest, and elements from the statistical theory of signal detection, and estimate the parameters of any particular mode. I showed in Chapter 3 that neutron data can appear to have splittings characteristic of the normal modes of the Sun, and I briefly used a combinatorial argument to justify the appearance of these splittings. Here, I am about to take a parametric approach to these composite spectra.

5.2 Signal Estimation and Detection

In the statistical theory of signal detection, which forms the underlying theory for e.g. radar, communications systems and array processing, the observations $x(t)$ are assumed to be generated by a probabilistic mechanism acting on the output of a source which can produce, say, 0’s and 1’s. The classic binary detection problem can
be phrased in terms of two hypotheses

\[ H_0 : x(t) = n(t); \]

\[ H_1 : x(t) = s(t) + n(t) \]  \hspace{1cm} (5.13)

where \( x(t) \) is the process in the observation space that is received, \( n(t) \) is a noise process, sometimes assumed to be Gaussian with a particular covariance structure, and \( s(t) \) is a random signal having particular properties. In case \( H_1 \) the source has generated a signal and in case \( H_0 \) it has not. Define the likelihood ratio,

\[ \Lambda(X) = \frac{p_{X|H_1}(X|H_1)}{p_{X|H_0}(X|H_0)} \]  \hspace{1cm} (5.14)

where \( p_{X|H_1}(X|H_1) \) is the probability distribution of the observation \( X \) when \( H_1 \) is true, and similarly for \( p_{X|H_0}(X|H_0) \). When the likelihood ratio (or the log likelihood ratio) exceeds a threshold \( \eta \) or \( \log \eta \), chosen according to a Neyman-Pearson or Bayesian criterion, one chooses \( H_1 \), and \( H_0 \) otherwise. This “classical” theory is well described in a number of excellent texts, namely Middleton [72], Van Trees [122], Davenport and Root [31] and Helstrom [53].

If the values of the parameters \( \omega_0, \alpha, \rho, \) and \( A \) were known, then the detection problem, that is, the decision between \( H_0 \) and \( H_1 \) could be solved using a matched filter (see [122] p. 591-597], and also §5.2). That is, one would correlate the data with the expected form of the solution to (5.3), substitute this into the likelihood ratio and base the decision on whether the likelihood ratio exceeds the threshold. Unfortunately, the problem at hand is more difficult. Since the parameter values are not known, one must solve both the detection and estimation problem at the same
The simultaneous detection and estimation problem can be done using an estimator-correlator approach, which is described in [54, §11.1.9]. To summarize, presupposing that a signal is present, one first uses a technique such as maximum likelihood to estimate the unknown parameters, and then one constructs a receiver for detecting a signal having these parameter values. The receiver takes the estimated parameters and generates a filter matched to the signal expected, filters the input (correlates or convolves the data with the filter) and the squared output at the end of the filtered input forms a test statistic for the test of $H_0$ vs $H_1$ above. When this test statistic exceeds a certain prespecified decision level, the decision is to reject $H_0$, i.e. the system decides that the signal is present. I will concentrate only on the estimation problem in this thesis for simplicity, though the detection problem is tractable, see §6.2.2. The estimation and detection portions ultimately both depend on an understanding of eigenvalue decompositions for stationary processes, or Karhunen-Loève decompositions.

5.3 Eigenvalue Representations and the Karhunen Loève decomposition

The Karhunen-Loève expansion of the autocorrelation function of the output signal of the system can be used to write the likelihood of the observations in terms of the parameters. The Karhunen-Loève expansion allows one to write a stationary process as a linear combination of the orthonormal eigenfunctions of its autocovariance sequence, which has the effect of producing uncorrelated coefficients, see Appendix C. For a Gaussian process, the likelihood function for the process is reduced to a product of individual Gaussian densities for the coefficients. Since autocovariance matrices for
5.3. EIGENVALUE REPRESENTATIONS

stationary processes are Toeplitz\(^4\) and positive definite, the eigenvalues are distinct and real. Let \( R(t, u) = R(t - u) = \text{Cov}\{x(t)x^*(u)\} \) be the autocovariance function for the stationary, zero-mean process \( x(t) \). Let the eigenfunctions \( \psi_n(t) \) and eigenvalues \( \theta_n \) be solutions to the eigenvalue problem

\[
\int_{-T}^{T} R(\tau - u) \psi(u) du = \theta \psi(\tau), \quad -T \leq \tau \leq T \tag{5.15}
\]

and one can form a series expansion of the process \( x(t) \) by expanding the continuous time signal \( x(t) \) in the eigenfunctions of the kernel

\[
x(t) = \sum_{n=0}^{N-1} c_n \psi_n(t), \quad \text{where} \quad c_n = \int_{0}^{T} x(t) \psi_n(t) dt \tag{5.16}
\]

for \( 0 \leq t \leq T \) and \( n = 0, \ldots, N-1 \) gives a vector \( c = [c_1 c_2 \ldots c_N]^T \) of eigencoefficients which can be used to rewrite the likelihood. The decomposition in (5.16) is referred to as the Karhunen-Loève expansion. Suppose the data are Gaussian distributed, and if complex, are circularly-symmetric complex Gaussian distributed\(^5\). The expansion of the signal on the orthogonal set of eigenfunctions \( \psi_n(t) \) results in independent coefficients \( c_n \), so that the (log) likelihood is a product of the individual Gaussian densities and can be written, in the complex case, as

\[
L(\{x\}) = \prod_{n=0}^{N-1} \frac{1}{\pi \theta_n} e^{-|c_n|^2/\theta_n}; \quad -\ell(\{x\}) = N \ln \pi + \sum_{n=0}^{N-1} \ln \theta_n + \frac{|c_n|^2}{\theta_n}. \tag{5.17}
\]

\(^4\)Toeplitz matrices are matrices \( R = [r_{ij}]_{i,j=0}^{N-1} \) for which the relation \( r_{ij} = r_{kl} \) if \( i - j = k - l \) holds for all \( 0 \leq i, j, k, l \leq N - 1 \).

\(^5\)Circularly symmetric complex time series \( x(t) \) have zero mean and the property that the \( \text{Cov}\{x(t)x(u)\} = 0 \), for all \( t, u \), where there is no complex conjugate on the second term. This discrepancy is important since complex demodulation on real processes will be used later on in this chapter. I will refer to such processes as complex normal.
5.3. EIGENVALUE REPRESENTATIONS

see [III] §XII], or in the real case, as

\[ L(\{x\}) = \prod_{n=0}^{N-1} \frac{1}{\sqrt{2\pi\theta_n}} e^{-1/2(c_n^2/\theta_n)}; \quad -\ell(\{x\}) = N \ln 2\pi + \sum_{n=0}^{N-1} \ln \theta_n + \frac{|c_n|^2}{2\theta_n}. \] (5.18)

and the covariance matrix for the \(c_n\)'s is the diagonal matrix of \(\theta\)'s, i.e. diag\(\{\theta_1, \theta_2, \ldots, \theta_N\}\).

Note that using Mercer’s theorem, see Appendix D.1, it is easy to show that \(E\{|c_n|^2/\theta_n\} = 1\), when \(x(t)\) has covariance function \(R(t,s)\) with eigenvalues and eigenfunctions as above.

Solving the Associated Eigenvalue Problem

Returning to the Lorentzian profile (5.4) used by Anderson, Duvall and Jefferies for the spectrum and taking the inverse Fourier transform, the autocovariance function of the Lorentzian signal is

\[ R_s(\tau) = Ae^{i\omega_0 \tau} e^{-\alpha |\tau|/2} \] (5.19)

the autocovariance functions under the null and alternative hypotheses are

\[ R_0(\tau) = \sigma^2 \delta(\tau) \] (5.20)

\[ R_1(\tau) = Ae^{i\omega_0 \tau} e^{-\alpha |\tau|/2} + \sigma^2 \delta(\tau) \] (5.21)

respectively, assuming stationary Gaussian white noise and a stationary Gaussian signal. The shift in frequency \(\omega - \omega_0\) in (5.4) can be interpreted as a complex de-modulate, that is, multiplication of the data sequence \(x(t)\) by \(e^{-i\omega_0 t}\) (see Appendix A.5.3). This autocovariance structure is part of a class of processes having rational

6Concisely, Mercer’s theorem states that \(R(s,t) = \sum \theta_n \psi_n(t)\psi_n(s)\) for all \(s,t\) with absolute, uniform convergence.
spectra for which the solutions $ψ_n(t), θ_n$ to the eigenvalue problem (5.15) can be determined analytically.\footnote{Incidentally, the eigenvalue problem in which the autocovariance kernel represents bandlimited white noise has been solved by Slepian [99, 101], and the solutions are the familiar prolate spheroidal functions.} Substituting the autocovariance of $x(t)$ under $H_1$, (5.21) into the eigenvalue problem (5.15), one finds a solution to the integral equation by solving the associated differential equation, see Kailath [59], Helstrom [53, p 136-141] or Van Trees [122, p 541-3]. Briefly, the integral equation above has solutions only when $0 < θ < \frac{4A}{α}$. The solutions can be written in terms of the angular frequencies $b$ which are solutions to

$$(\tan bT + \frac{2b}{α})(\tan bT - \frac{α}{2b}) = 0$$

ordered $0 = b_0 < b_1 < b_2 < \ldots$, where $b_n$’s with even $n$ are associated with the zeros of the first term of (5.22) and the odd $b_n$’s with zeros of the second term. The $b_n$’s can be thought of as the locations $b_n$ where the functions $-2b/α$ or $α/2b$ intersect with $\tan bT$. Fig. 5.2(a) gives an example of graphically solving for $b_n$’s is shown where parameter values $A = 1$, $α = 1/2$, $T = 2$ are used.

The solutions to the eigenvalue problem (5.15) in terms of the $b_n$’s, are then

$$θ_n = \frac{Aα}{b_n^2 + (α/2)^2} + σ^2$$

$$= \frac{Aα}{b_n^2 + (α/2)^2} + \frac{A}{ρ}; \ n = 1, 2, \ldots$$

where $ρ = A/σ^2$ is the signal to noise ratio when $σ^2 \neq 0$. The fact that the Lorentzian eigenvalues $θ_n$ are exactly equal to the spectrum at the frequencies $b_n - ω_0$ is not true in general.\footnote{For stationary processes however, Szegő’s theorem, see Appendix D.2 implies that the eigenvalues and the spectrum are equal in a certain sense.} From this perspective, $b_n$ can be thought of as a sort of angular frequency
5.3. EIGENVALUE REPRESENTATIONS

Figure 5.2: In panel (a), the solutions $b_n$ to equation (5.22) are the intersections of $\tan bT$ and $-2bT/\alpha T$ (long-dashed line, even $n$) or $\alpha T/2bT$ (dot-dashed curve, odd $n$). In both figures $\alpha T/2$ is set to 1, and in panel (b) $T$ is set to 2. Panel (b) shows the eigenfunctions $\psi_n(t)$ (eigenfunctions are defined on $[-T, T]$, but the negative half is not shown). Approximately sinusoidal curves represent $e^{-i\omega_0 t}\psi_n(t)$’s with even $n$, where the number of nodes increases with $n$. Cosinusoidal curves represent $e^{-i\omega_0 t}\psi_n(t)$’s with odd $n$.

separation of a grid of unequally spaced frequencies from 0 to $\pi$. If the $b_n$’s are sorted in ascending order, the $\theta$’s will be sorted in descending order, $\theta_0 > \theta_1 > \ldots$. This gives the eigenfunctions

$$
\psi_n(t) = \begin{cases} 
    e^{i\omega_0 t} \cdot (\cos b_n t) \cdot \left(1 + \frac{\sin 2b_n \pi}{2b_n T}\right)^{-1/2} \cdot T^{-1/2} & n \text{ odd, } -T \leq t \leq T \\
    e^{i\omega_0 t} \cdot (\sin b_n t) \cdot \left(1 - \frac{\sin 2b_n \pi}{2b_n T}\right)^{-1/2} \cdot T^{-1/2} & n \text{ even, } -T \leq t \leq T 
\end{cases}
$$

which are sinusoidal with period $1/b_n$. Fig. 5.2(a) shows the eigenfunctions associated with the solutions $b_n$ to (5.22). Substitution of these eigenfunctions and eigenvalues into (5.16) and then into (5.17) gives the value of the likelihood for known values of the parameters $\omega_0, \alpha, \rho,$ and $A$. In practice, the $b_n$’s are approximated with a
5.4 Parameter Estimation

To solve the estimation problem, one can numerically minimize the likelihood. Fortunately, expressions for the Karhunen-Loève eigenfunctions and eigenvalues of the Lorentzian are known, so the gradient of the likelihood is needed for this minimization.

5.4.1 Partial Derivatives

Maximization of the likelihood in (5.17) is equivalent to minimizing the negative log-likelihood (complex case)

\[ -\ell = \sum_{n=0}^{N-1} \ln(\pi \theta_n) + \frac{|c_n|^2}{\theta_n}. \quad (5.27) \]

where \( \theta_n \) are dependent upon \( A, \alpha, \) and \( \rho \), the \( c_n \)'s depend on parameters \( \omega_0 \) and \( \alpha \), and meanwhile the \( b_n \)'s depend only on \( \alpha \), where \( T \) is fixed. Surprisingly, it is possible to obtain the gradient needed to minimize the negative log likelihood. The derivative
of \((-\ell)\) with respect to \(A\) is, substituting the equation for the \(\theta_n\)'s, \((5.23)\),

\[
\frac{d}{dA}(-\ell) = \frac{1}{A} \left( N - \sum_{n=0}^{N-1} \frac{|c_n|^2}{\theta_n} \right). \tag{5.28}
\]

The derivative of \((-\ell)\) with respect to \(\rho\) is easily found to be

\[
\frac{d}{d\rho}(-\ell) = \frac{A}{\rho^2} \sum_{n=0}^{N-1} \frac{1}{\theta_n} \left( 1 - \frac{|c_n|^2}{\theta_n} \right). \tag{5.29}
\]

The derivative of \(|c_n|^2\) with respect to \(\omega_0\) is

\[
\frac{d}{d\omega_0}|c_n|^2 = i \int_{-T}^{T} \int_{-T}^{T} (t-u)\psi_n(t)\psi_n^*(u)x(t)x^*(u)dtdu \tag{5.30}
\]

so that the derivative of \(-\ell\) with respect to \(\omega_0\) is

\[
\frac{d}{d\omega_0}(-\ell) = \sum_n \frac{1}{\theta_n} \frac{d}{d\omega_0}|c_n|^2. \tag{5.31}
\]

from \((5.25)\) and \((5.16)\). Finally, the derivative of \((-\ell)\) with respect to \(\alpha\) must be done carefully, as the coefficients \(b_n\) are a function of \(\alpha\), \((5.22)\), and these are used both in the expression for \(\theta_n\) and \(c_n\). To find the derivative of \(b_n\) with respect to \(\alpha\), differentiate the first and second terms in Eqn. \((5.22)\) implicitly

\[
\frac{db_n}{d\alpha} = \begin{cases} 
\frac{1}{2b_n}(T \sec^2 b_n T + \frac{\alpha}{2b_n^2})^{-1} & \text{n odd}, \\
\frac{2b_n}{\alpha^2}(T \sec^2 b_n T + \frac{2}{\alpha})^{-1} & \text{n even, or}
\end{cases}
\]

\[
= \frac{2b_n}{T\alpha^2 + 4Tb_n^2 + 2\alpha} \quad \text{n even or odd.} \tag{5.33}
\]
where the identity \( \sec^2 b_n T = 1 + \tan^2 b_n T \) and the relation (5.22) was used. Now the derivative of \( \theta_n \), Eqn. (5.23), with respect to \( \alpha \) is

\[
\frac{d\theta_n}{d\alpha} = A \cdot \frac{b_n^2 - (\alpha/2)^2 - 2\alpha b_n \frac{db_n}{d\alpha} + \alpha^2/2}{(b_n^2 + (\alpha/2)^2)^2} \tag{5.34}
\]

while the derivative of \( \psi_n(t) \), Eqn. (5.25), with respect to \( \alpha \) is

\[
\frac{d\psi_n(t)}{d\alpha} = \begin{cases} 
\psi_n(t) \cdot \frac{db_n}{d\alpha} \cdot \left[-t \tan b_n t - \frac{2b_n T \cos 2b_n T - \sin 2b_n T}{2b_n (2b_n T + \sin 2b_n T)}\right] & \text{n odd;} \\
\psi_n(t) \cdot \frac{db_n}{d\alpha} \cdot \left[t \cot b_n t + \frac{2b_n T \cos 2b_n T - \sin 2b_n T}{2b_n (2b_n T - \sin 2b_n T)}\right] & \text{n even.} 
\end{cases} \tag{5.36}
\]

Taking the derivative of the expression for \( |c_n|^2 \) and using the product rule gives

\[
\frac{d}{d\alpha} |c_n|^2 = \int_T^{-T} \int_T^{-T} \left[ \frac{d}{d\alpha} \left[ \psi_n(t) \psi_n^*(u) + \psi_n(t) \frac{d}{d\alpha} \left[ \psi_n^*(u) \right] \right] x(t)x^*(u) dt du \right] \tag{5.37}
\]

where the first expression is for \( n \) odd and the second is for \( n \) even. The final expression for \( \frac{d}{d\alpha} (\cdot) \) is then

\[
\frac{d}{d\alpha} (\cdot) = \sum_{n=0}^{N-1} \frac{1}{\theta_n} \cdot \left[ \frac{d}{d\alpha} |c_n|^2 + \left( 1 - \frac{|c_n|^2}{\theta_n} \right) \cdot \frac{d}{d\alpha} \left[ \theta_n \right] \right] \tag{5.38}
\]

where the derivative Eqns (5.34) and (5.37) as well as the definitions (5.24) and (5.16) with (5.25) can be substituted to find \( \frac{d}{d\alpha} \). Note that most of the derivatives vanish when \( |c_n|^2 \) takes the value \( \theta_n \) which is consistent with the fact that \( \mathbb{E}\{|c_n|^2\} = \theta_n \) when \( x(t) \) has a Lorentzian spectrum.
5.4.2 Advanced Optimization

The minimization of the log-likelihood (5.27) can be done numerically using an advanced optimization routine. Beginning with reasonable estimates for the center frequency $\omega_0$, amplitude $A$, SNR $\rho$, and line width $\alpha$ (usually estimated by measuring the linewidth in radians at the half power point and then adjusted by subtracting the width of the spectral window, approximately $2W$ for a multitaper spectrum estimate), one can use an iterative procedure which uses the derivatives given above to eventually converge to the minimum of the function.\(^9\) Determining whether or not the likelihood is convex is nontrivial, and it is discussed in §6.2.

5.5 Practical Considerations for Real Data

As mentioned earlier in the chapter, the goal of this analysis is to estimate mode parameters in data which contain more than one mode. In the case of helioseismic data such as the GONG data considered in §3.1, spectra for each $l,m$ will contain well-spaced peaks for each $n$. In the case of the neutron monitor spectra, modes in the $p$–mode band and their splittings will be superimposed, so the difficulty of isolating each individual mode is much greater.

Fortunately, one can isolate bands in frequency using complex demodulation, which is discussed in Appendix A.5.3. The result of complex demodulation of a time series $x(t)$ is to shift it in frequency by the angle $\omega_1$ (equivalently, multiply by $e^{-i\omega_1 t}$), apply an appropriate lowpass filter with bandwidth $W$, and then downsample the series so that $\delta t = 1/(2W)$. When one uses complex demodulation, the result

\(^9\)A quadratic trust-region algorithm fminunc is implemented in the MATLAB (Natick, MA) Optimization toolbox, and I have used fminunc for the examples in this chapter. See MATLAB documentation for implementation details.
is a complex sequence \( x \downarrow (j; f) = \hat{x}(t_j; f) \) with significantly fewer samples than the original time series. The spectrum of the series is a portion of the spectrum of the original series, shifted to center at 0, and is not in general symmetric about 0, which means the associated time series is complex.

The eigenvalues \( \theta_n \) and eigencoefficients \( c_n \) of the complex demodulate of the series are computed in the same way as that of the original process (except the \( \psi_n(t) \)'s used are downsampled), and one discards the \( \theta_n \)'s and \( c_n \)'s which correspond to \( b_n > 2\pi W \), where \( W \) is the bandstop of the filter in Hz. One can then fit the parameters by minimizing the same likelihood as before, but omitting terms with higher \( n \).

5.6 Examples

Some simple simulations are used to illustrate the estimation problem in this section.

5.6.1 Lorentzian Likelihood - with and without noise

Beginning with a simple example, simulate a process with Lorentzian spectrum having the parameters \( f_0 = 0.2 \text{Hz} \) (\( \omega_0 = 1.26 \text{rad/sec} \)), \( \alpha = 0.0013 \) (\( Q = 1000 \)), \( A = 10 \) and let the additive noise power be \( \sigma^2 = 1 \). Generate a process having \( N = 256 \) samples by generating random phase for the Lorentzian, and inverse Fourier transforming the result.

Solving the equations for \( b_n \) \((5.22)\) using Newton’s method, and then substituting these into \((5.23)\) to get the \( \theta_n \)'s and into \((5.25)\) to get the \( \psi_n(t) \)'s (where \( \psi_n(t) \) is valid on \([-T, T]\)) expand the simulated process on the basis \( \{\psi_n(t)\}_n \) to get \( c_n \). Figure 5.3 shows \(|c_n|^2\) and \( \theta_n \) for two such simulations, both with and without additive noise. Recall that the values of \( \theta_n \) are equal to the values of the spectrum at the
Figure 5.3: Panel (a) shows a realization of a Lorentzian process eigencoefficients $|c_n|^2$, with no additive noise and panel (b) shows the Lorentzian spectrum in additive noise with power $\sigma^2 = 1$. The eigenvalues for both processes are shown by the smooth black curve and are identical to the spectrum.

angular frequencies $b_n$. When additive noise is present, the eigenvalues are adjusted to $\theta_n + \sigma^2$. The negative log likelihood (real version) evaluates to $-\ell = 105.6$ for the eigencoefficients in panel (a), no noise, and $-\ell = 339.3$ for the eigencoefficients in panel (b), additive noise with power $\sigma^2 = 1$.

**Minimizing $(-\ell)$**

Upon substitution of the process whose eigencoefficients and eigenvalues are shown in Fig. 5.3(b) for the true values of the parameters into an advanced optimization algorithm for minimizing $-\ell$, initialized away from the true values of the parameters, one obtains the parameter estimates $A = 9.73$, $\rho = 10.1$, $\alpha = 0.00144$, $\omega_0 = 1.26\text{rad/sec}$ ($f_0 = 0.201\text{Hz}$), which are reasonably close to the true values.
Figure 5.4: In panel (a), the upper axis shows the magnitude of the complex demodulate (thick blue curve) of the process shown in Fig 5.4 (b) along with the downsampled magnitude (open circles) and the original process (in grey), where the complex demodulate is centered at the frequency $f_1 = 0.2$Hz and the bandwidth $W = 0.0625$Hz. The phase of the complex demodulate (lower axis) is relatively constant since the complex demodulate is centered at the peak frequency, but note it is wrapped to 180°. In (b), the magnitude-squared eigencoefficients (in blue) are plotted against the eigenvalues, $\theta_n$ (smooth black curve). The complex demodulate retains the peak shape.

**Complex Demodulates**

Because the values of $\theta_n$ and $|c_n|^2$ are really only of interest near the peak, it suffices to do complex demodulation on the data before doing the eigendecomposition and discard $|c_n|^2$ and $\theta_n$ which correspond to frequencies $b_n$ outside the filter passband.

Figure 5.4 shows the result of complex demodulation of the series with additive noise, where the series was decimated by a factor of $16 = 1/(2W)$, so the new Nyquist frequency is 0.0625Hz. The appearance of the magnitude of the complex demodulate is like the envelope of the signal, while the phase appears to be reasonably constant near 170°. If one were to compute the complex demodulate at a frequency $\omega_1$ away
from $\omega_0$ one would find the phase had a slope proportional to $\omega_0 - \omega_1$. In any case, expanding the complex demodulate on the first 16 eigenfunctions results in the eigenvalues and eigencoefficients in Fig. 5.4(b). The $|c_n|^2$s that result are faithful to the Lorentzian peak shape. The $|c_n|^2$s and $\theta_n$'s for the demodulated process can be substituted into the negative log-likelihood and minimized in much the same way as the original process.
5.6.2 Seismic Data

To illustrate the methods given in this chapter, I give seismic data in lieu of helioseismic data. The reason for this is that the Lorentzian model has long been used in the seismology literature, and comparatively more is known about the Earth’s normal modes of oscillation than those of the Sun. The most important reason, however, is that the quality factors of the Earth’s normal modes is generally near 200-300, while the helioseismic modes generally have Q of several thousand. When only short data series are available, one may lack the resolution to adequately distinguish high-Q modes from sinusoids. Therefore in this section, seismic data serves to motivate better these results. That said, a short introduction to seismic data analysis is in order.

The free oscillations of the Earth can be excited by earthquakes and other anthropogenic sources such as nuclear testing, and these can start the Earth ringing like a bell for several days.
5.7 Summary

In this chapter, I gave a literature review of mode-fitting methods, most of which were based on the spectrum, with the exception of the methods of Schou. Taking a different approach on a simple Lorentzian model, it was possible to analytically find the Karhunen-Loève eigenfunctions and eigenvalues for processes having Lorentzian autocovariance.

Two problems were then identified: the (ii) detection of peaks with Lorentzian structure, and (i) the estimation of these peaks’ parameters. First, maximum-likelihood estimation could be carried out by substituting the Karhunen-Loève expansion directly into the process likelihood, assuming Gaussian statistics. Because of orthogonality of the Karhunen-Loève eigenfunctions, this simplified the covariance so the likelihood could be re-written as a product of independent Gaussian densities. Using a numerical optimization routine, one can then minimize the negative log-likelihood over the parameter space.

Using results from communication theory, it is possible to solve the detection problem (ii) using a simplified statistic, $U$, based on the likelihood ratio. The decision between $H_0$ and $H_1$ can be done by comparing the test statistic $U$ to a threshold value $U_0$ determined from a prespecified false detection rate. If $U > U_0$ then the signal is determined to be present. The detection problem is left for future work, as the implementation is complicated by a number of additional factors, described in §6.2.2.

Finally, some simple simulations were carried out, with some important adaptations, namely (i) doing Karhunen-Loève eigenexpansions on complex demodulates and (ii) estimating parameters from processes having transfer function (5.3). The advantage of this approach to mode fitting, as compared to the mode fitting methodologies
mentioned in section 5.1 is that it is a (a) simple, (b) appropriate for composite spectra because the estimation procedure can be done on complex demodulates of the data, and (c) there is a closely-related detection procedure for determining if the estimated mode is actually present with a certain significance.
Chapter 6

Summary and Conclusions

6.1 Summary

In Chapter 2 I gave an overview of solar physics and helioseismology. The main gist of the chapter was that the study of helioseismology has given a lot of information about the structure of the solar interior - for example the location of the solar tachocline, whose properties are largely associated with the solar magnetic dynamo. The discovery of modal structure in the interplanetary magnetic field was astounding to many because it was believed that coherent modal structure is destroyed by the same vigorous convection that excites the modes in the first place.

In Chapter 3 I investigated a number of data sources for mode structure, beginning with GONG helioseismic time series, in which I calculated an \((l, \nu)\) diagram which showed the structure of the oscillations on the isolated disk images spatially (through \(l\)) and temporally (through frequency \(\nu\)). The next example was the proton density data at ACE which monitors charged particle flux in the interplanetary magnetic field and I showed evidence that the distribution of the spectra of these data indicated an overwhelming modal component in the \(p\)-mode band. I then introduced the
possibility of solar modal structure in cosmic radiation on Earth, and gave examples of Bartol neutron data in the $p$–mode band with up to 29\% modal structure by power (see “Thule” column of Table 3.3.1). This discovery is entirely new.

Chapter 4 contained the main result of the thesis, namely the quantification of upcrossings of a given significance level expected in spectrum estimates of white noise. The result is effectively the false detection rate for periodic components in noise, a question which was raised by Schuster in 1898 and has remained unanswered until just now. Additional combinatorial arguments show that multiple comparisons problems are more prickly for spectrum estimates than in conventional statistical problems because the number of hypotheses to be tested is unknown, but our result on the dwell bandwidth of the peaks offers a combinatorial reason for the difficulty of such problems. It was also shown that the shape of a spurious peak in a multitaper estimate is roughly triangular, while genuine line components take the shape of the spectral window, which is roughly rectangular, by virtue of convolution with a delta-function when the data are perfectly sinusoidal.

The theoretical results of Chapter 4 were applied to the ACE proton density data and the neutron monitor data in the form of a permutation test. Spectra were computed on randomly permuted versions of these time series, and it was found that they contained several times more peaks than expected at high significance levels. This indicated a change in distribution from the null hypothesis - more significant peaks than expected could be the result of the contributions of large numbers of modes.

In Chapter 5, a parametric approach, borrowed from the communications literature, was adopted to characterize individual modes in helioseismic spectra. I began
6.2. FUTURE WORK

with a brief overview of the methods used in the literature for estimating the helioseismic mode parameters: amplitude $A$, damping $\alpha$, or $Q$, center frequency $\omega_0$, and signal to noise ratio, $\rho$, and then introduced a maximum likelihood method for estimation and detection of individual modes. Complex demodulation was necessary to avoid including contributions from far-field signals.

6.2 Future Work

6.2.1 Part I: Nonparametric Methods

MSC crossing rate

As an extension to Chapter 4, one might wish to find the upcrossing rate of the magnitude squared coherence (MSC estimation is discussed in Appendix A). Given that MSC estimates are routinely computed for the purpose of determining periodicities common to two series at once, the question of how many spurious peaks one obtains is also of interest here. Specifically, when the two series actually have zero coherence, how many large peaks ought there to be in an MSC estimate?

Unfortunately, this problem may be much more difficult than the $F$–test to solve. The result of an “asymptotic” MSC crossing rate simulation similar to the one computed in §4.4.1, Fig. 4.4 is shown in Fig. 6.1. The overall impression from Fig. 6.1 is that with more degrees of freedom and a larger number of samples, statistically fewer spurious crossings of any given level will result, but it is difficult to make any more meaningful conclusions about the particular relation between significance level and crossing rate.
6.2. FUTURE WORK

Figure 6.1: Expected crossing rates of the magnitude squared coherence computed using two random normally distributed sequences with zero coherence. For reference, the solid black curve marked “P” is the expected crossing rate of the periodogram for any given level of significance, and the dotted curves represent the result of simulations using $W = 1/1000$ and letting $N = 1,000$ (red, marked $NW = 1$), $N = 5,000$ (orange, marked $NW = 5$), $N = 10,000$ (green, marked $NW = 10$), and $N = 100,000$ (blue, marked $NW = 100$).

6.2.2 Part II: Parametric Methods

Bandwidth of Complex Demodulate

The Lorentzian spectrum has an effective bandwidth of $\alpha T/2$. One can also measure the bandwidth by twice the frequency shift of the spectrum $-3$dB (or half-power point) away from the maximum. For composite spectra, the methods given in Chapter 5 must be done on complex demodulates of the data, so the question becomes - what is the optimum bandwidth $W$ for the complex demodulate given a value for the Lorentzian bandwidth?
6.2. FUTURE WORK

Detection of Lorentzian Profile in Noise

So far in this thesis, I have only attempted the estimation problem for Lorentzians in noise, but it appears the detection problem is also tractable. The detection problem can be done by way of an appropriate matched filter (see [122]) now that the parameters have been estimated; the test statistic is the $T$th output of the filtered version of $x(t)$. In fact, when a matched filter is used to detect a signal with Lorentzian spectrum, much is known about the (false) detection probability as a function of the signal to noise ratio, especially when both signal and noise are Gaussian and stationary, and the noise is white. I will leave out the details of the design of the matched filter, but see Helstrom [54] p. 304-305, Ch 11 and Appendix H for a derivation.

The detection problem can be done by comparing the test statistic

$$U = \frac{1}{2\sigma^2} \sum_{n=0}^{N-1} \frac{\lambda_n|c_n|^2}{1 + \lambda_n}$$  \hspace{1cm} (6.1)

where $\lambda_n$ is related to $\theta_n$ and is defined below, with a pre-determined decision level $U_0$, and $U_0$ is chosen to satisfy

$$q_0(U_0) = \Pr\{U > U_0|H_0\} = \int_{U_0}^{\infty} p_0(u)du$$  \hspace{1cm} (6.2)

where $q_0$ is the desired false alarm probability for the test and $p_0(u)$ is the probability density of $U$ under hypothesis $H_0$. $U$ is derived from the log of the likelihood ratio (5.14) for this problem, discarding a known scale factor.

The probability density of $U$ is often derived first by calculating the characteristic function (or moment generating function) of $U$ and then inverse Fourier transforming. This calculation is prohibitively difficult except in the case where the signal and noise
are both Gaussian processes and are independent of one another.

Letting \( \lambda_n \) be the \( n \)th eigenvalue of the kernel \( \sigma^{-2}R_s(\tau) \) (5.19), that is

\[
\lambda_n \psi_n(t) = \frac{1}{\sigma^2} \int_0^T R_s(t, \tau) \psi_n(\tau) d\tau, \quad 0 \leq t \leq T
\]  

(6.3)

where \( \lambda_n \) is related to \( \theta_n \) by \( \lambda_n = \theta_n/\sigma^2 - 1 \) and the eigenfunctions \( \psi_n \) are the same as the eigenfunctions of 5.21\(^1\).

The false-alarm probability when the cutoff is \( U_0 \) can be shown to be

\[
q_0(U_0) = \sum_n \frac{r_n D(1)}{(1 + \lambda_n)} \exp \left\{ -\frac{1 + \lambda_n}{\lambda_n} U_0 \right\}
\]  

(6.4)

where the quantity \( r_n \) is defined as

\[
r_n = \frac{(-1)^n 4 \lambda_n (K - \lambda_n)}{K (mK + 2 \lambda_n)} e^m,
\]  

(6.5)

the Lorentzian effective bandwidth is \( m = \alpha T/2 \), \( K = A/(\sigma^2 \alpha) = \rho/\alpha \), and the Fredholm determinant, \( D(z) \), is

\[
D(z) = \frac{(g + 1)^2 e^{mg} - (g - 1)^2 e^{-mg}}{4g} e^{-m}
\]  

(6.6)

where \( g = \sqrt{1 + Kz} \). \( U_0 \) is chosen so that the test has a tolerable false alarm rate. Because the \( \lambda_n \)'s can be large for large \( \alpha \), a great number of terms must be included for convergence of this sum close to the true value. The detection probability of the

\(^1\)This is because any complete orthonormal set can be eigenfunctions for white noise (with all eigenvalues equal to 1 when the noise has been normalized to unit variance).
Lorentzian signal in white Gaussian noise can be written as

\[ q_1(U_0) = \sum_n r_n \exp \left\{ -\frac{U_0}{\lambda_n} \right\} \quad (6.7) \]

and it depends upon the energy-to-noise ratio, \( AT/\sigma^2 = \rho T \) as well as the Lorentzian bandwidth, \( m = \alpha T/2 \). Detection probabilities are plotted in Fig. (11-2) of Helstrom [54] for various choices of ENR and effective bandwidth. For example, when \( q_0 \) is set to \( 10^{-6} \), the effective bandwidth \( m = 20 \) and the average ENR is 49, one has a better than 0.8 probability of detecting the signal. This seemingly small false-detection tolerance is typical for communications problems, and depending on the application, is often set much smaller.

In the future, I expect to use this detection paradigm on real data. However, some practical details arise, namely (i) the use of the detection statistic on complex demodulates of the data is more difficult, and (ii) the inversion of equation (6.4) requires advanced numerical methods. These two difficulties combined, have prevented me from including the detection problem in the present work.

Another yet-unsolved question lies in the convexity properties of the likelihood. There are few examples of convex likelihoods in the literature, for practical comments on score functions of Gaussian processes, see e.g. [103].

**Frequency Modulation**

I’ve mentioned the tendency of the modal frequencies to change with time, perhaps with a \( \sim 1.2 \) year period [57]. This can be modeled if center frequency of the mode \( \omega_0 \) is allowed to vary with time, i.e. \( \omega_0 = \omega_0(t) \). Letting the complex demodulate of the process \( x(t) \) at frequency \( \omega_1 \) be denoted \( x_{\omega_1}(t) \), the autocovariance function of
\( x_{\omega_1}(t) \) is
\[
R_{\omega_1}(\tau) = e^{-i\omega_1 \tau} R_x(\tau) = Ae^{i(\omega_0 - \omega_1) \tau} e^{-\alpha |\tau|/2} + \sigma^2 \delta(\tau) e^{-i\omega_1 t}.
\] (6.8)
evaluating the phase of the autocorrelation at 1, or equivalently, estimating the phase of the complex AR(1) coefficient of the equivalent AR model results in
\[
\text{Im}\{\log R_{\omega_1}(1)\} = (\omega_1 - \omega_0)
\] (6.9)
which reveals the frequency separation between the frequency \( \omega_0 \) and \( \omega_1 \). This technique could be used along with the model described in Chapter 5 to produce estimates of the changes in phase of the complex model. For an interesting approach to estimating polynomial phase, see [115]. The Cramér-Rao bound for estimating polynomial frequency modulation is given in [82, 83].

Application of frequency modulation estimation to helioseismic data could lead to better error estimates of the modal frequencies at different points in the solar cycle. I suspect FM may be the cause for bad reproducibility and inconsistencies between mode frequency estimates from different temporal epochs and across different data products.

### 6.2.3 Future Data Analysis

**Muon Data**

Muon observations are complimentary to neutron monitor observations, as the neutron monitor observations extend from the lowest energies accessible to ground based observation up to approximately 50GeV, and surface muon observations have significant responses from 10GeV to several hundred GeV [85].
Figure 6.2: The local fractional variance spectrum for the multidirectional muon Data is shown. Data are divided into 4 year blocks and years on the plot refer to the beginning of each block. The darker bands near 1980 occur around the time of the Hydro Quebec failure. The interesting feature of this spectrum is that even though the data blocks are spaced with 66% overlap, one sees ridges that persist for longer than 3 blocks vertically and appear to “wander” with time. This artifact may be real, or may lack a physical basis.

The Nagoya multidirectional muon telescope gives hourly observations of cosmic ray intensity beginning in October 1970 to the present. These data are available publicly from [http://www.stelab.nagoya-u.ac.jp/ste-www1/div3/muon/muon1.html](http://www.stelab.nagoya-u.ac.jp/ste-www1/div3/muon/muon1.html). Multidirectional muon telescopes can detect anisotropy in the incoming secondary cosmic radiation, which can give directional information about geomagnetic storms, solar conditions and space weather.

Due to the multivariate nature of this dataset, the availability of only hourly data
(only $g$--modes can potentially be detected, with contributions by aliased $p$--modes), and the lack of any theoretical result in this thesis concerning the multivariate case, I have not included the analysis here. Due to the massive amounts of data and high reliability of the relative intensity measurements, it may be possible to track modal FM over time using the singular value decomposition paired with evolutive spectra.

An evolutive local fractional variance spectrum, see e.g. Mann and Park [69], of the muon data is shown in Fig. 6.2, where all 17 directions of data are used and data are combined into four-year blocks, overlapped by 66%. Two of the Delache-Scherrer $g$--mode candidates, [32], are highlighted in bands 2NW Rayleighs wide. The appearance of the band near 73.84uHz does not appear to persist at the same frequency for long, but instead may “wander” back and forth across the bandwidth, and may quickly “switch on” and off. The blue line shows the approximate year at which the data for g-mode candidates was measured. Note the large values of the spectrum between 1990 and 1995, this corresponds to the epoch near 1989 in which massive power failures occurred as a result of severe solar geomagnetic storms [16].

Investigation of muon data for $g$--mode frequencies could complement the results from the neutron data. As the muon measurements represent particles with greater energy, one could determine the extent of the solar effects on the cosmic rays as the energies increase. However, due to the hourly sampling of the data, the effect of aliased $p$--modes may compromise the detectability of $g$--modes.
6.3 Conclusion

Natural processes containing solar oscillations give interesting examples of *data that break the rules* (non-Gaussian, non-stationary, large numbers of lines, etc.). However, the solar example is a prominent one; because solar oscillation signatures have been identified in the interplanetary magnetic field, magnetosphere, ionosphere, atmospheric pressure, communications data (specifically dropped cell phone calls) as well as in the variations in cosmic radiation observed on Earth, they might contribute significantly to almost every other natural process observed in our solar system. From this perspective, it is important to develop methods suitable for time series containing many lines, as the implications could have important, constructive applications in science and engineering.


*Fourier Analysis of Time Series: An Introduction*.  


[69] M. E. Mann and J. Park. Oscillatory spatiotemporal signal detection in climate


Appendix A

Nonparametric Spectrum Estimation

The power spectrum of a continuous time stationary process is the Fourier transform of the autocorrelation function of the process. Thus it contains information about the temporal correlations in the data as a function of frequency. Delta functions in the power spectrum correspond to sinusoidal components in the data, and spectrum estimates are often computed for the purpose of determining where sinusoidal components contribute significantly to the total signal power. General overviews of spectrum analysis are available in [19, 17]. We give here a more concise summary of nonparametric spectrum estimation than in [48] beginning with direct estimators, the Welch method, and multitaper methods.

A.1 Direct Spectrum Estimators

Definition

Techniques for nonparametric estimation of power spectra began with the introduction of the periodogram by Schuster [96]. The periodogram is part of a larger class of direct spectrum estimators, which are computed as follows. Given equally-spaced
samples of a stationary, zero-mean time series \( x_t, t = 0, \ldots, N-1 \), the direct spectrum is

\[
\hat{S}^D(f) = \left| \sum_{t=0}^{N-1} x_t D_t e^{-i2\pi ft} \right|^2
\]  (A.1)

where \( D_t \) is a data window, often chosen to be one of those compared in [50] or [106] §2.6. The periodogram estimator results when \( D_t = \sqrt{1/N} \), \( t = 0, 1, \ldots, N-1 \) is substituted above. The nonuniform data window was introduced to reduce bias inherent in the periodogram. When a data window is used, \( E\{\hat{S}^D(f)\} = |D(f)|^2 * S(f) \), where \( D(f) \) is the Fourier transform of the data window \( D_t \), \( S(f) \) is the true power spectrum, and * denotes convolution.

Direct estimators are not only biased, but inconsistent, i.e. the variance of the estimator does not decrease as the number of samples is increased. The variance of this estimator can be reduced using Welch’s method, that is, by dividing the data into overlapped sections, computing a direct spectrum on each section and averaging the estimates. For more information on the Welch technique, see [127]. It can be shown that the Welch estimator is consistent when the block length is fixed. The Welch method is often applied to multitaper spectra, described in §A.2.

**Distribution of Direct Spectrum Estimators**

It can be shown that windowed Fourier transformed data is approximately complex-Gaussian distributed [68, 19] and is certainly asymptotically so using a weak central limit argument. As such, the direct spectrum at a given frequency is \( \chi^2_2 \) distributed, where uncorrelated real and imaginary parts each contribute a degree of freedom to the estimate, except at zero and the Nyquist frequencies, where the spectrum is real.
A.2 Multitaper Spectrum Estimators

Slepian Sequences

The multitaper method makes use of the discrete prolate spheroidal, dpss or Slepian sequences, \( \{ v_n^{(k)}(N, W) \}_{k,n=0}^{N-1} \), which form a set of finite-length, orthogonal sequences \( L^2 \)-optimally concentrated on the band \((-W, W)\) in frequency \[100\]. They satisfy the eigenvalue equation

\[
\sum_{m=0}^{N-1} \frac{\sin 2\pi W(n-m)}{\pi(n-m)} v_n^{(k)}(N, W) = \lambda_k(N, W) \cdot v_n^{(k)}(N, W)
\]  

(A.2)

where the eigenvalue \( \lambda_k(N, W) \) also denotes the fraction of energy in the band \((-W, W)\), and the index \( k \) sorts the sequences in descending order of energy concentration, \( 1 > \lambda_0 > \lambda_1 > \ldots > \lambda_{N-1} > 0 \). See Slepian, \[100\], for the rigorous formulation of the eigenvalue problem in discrete time, particularly the properties of the kernel which makes possible only a countable number of eigenvalues and eigensequences and the simultaneous timelimitedness and optimally-bandlimited properties of the Slepian sequences in frequency.\(^1\) The first \( K \approx 2NW \) of these are close to one, while the others rapidly drop to zero. Slepian sequences are best computed as the eigenvectors of a tridiagonal matrix, see \[112\].

\(^1\)The explicit dependence on \((N, W)\) will be dropped for notational convenience.
Multitaper Spectrum Definition

The multitaper estimator, \[111\], was developed as an approximate solution to the integral equation

\[
y(f) = \sum_{t=0}^{N-1} x_t e^{-i2\pi (t - \frac{N-1}{2})} = \int_{-1/2}^{1/2} \frac{\sin N\pi (f - \nu)}{\sin \pi (f - \nu)} dZ(\nu) \tag{A.3}
\]

obtained when the Fourier transform of the time series, \(y(f)\), is substituted into the Cramér spectral representation theorem, \[29\ 84\], namely

\[
x_t = \int_{-1/2}^{1/2} e^{i2\pi ft} dZ(f) \tag{A.4}
\]

where \(dZ(f)\) is a stationary, orthogonal increments process. When it exists, the spectrum of the process is then defined as \(S(f) = \text{E}\{|dZ(f)|^2\}\).

The multitaper spectrum is constructed as the local least-squares solution to (A.3) in the band \((f-W, f+W)\), and as such, is a weighted average of magnitude squared Fourier transformed tapered data sequences. Denoting the \textit{eigencoefficients} as

\[
y^{(k)}(f) = \sum_{t=0}^{N-1} x_t v_t^{(k)} e^{-i2\pi ft}, \tag{A.5}
\]

and the \textit{eigenspectra} as

\[
\hat{S}^{(k)}(f) = |y^{(k)}(f)|^2, \tag{A.6}
\]

the multitaper spectrum estimate is constructed as the following weighted average

\[
\hat{S}(f) = \frac{\sum_{k=0}^{K-1} d_k^2(f)\hat{S}^{(k)}(f)}{\sum_{k=0}^{K-1} d_k^2(f)} \tag{A.7}
\]
where the optimum frequency-dependent weights, $d_k(f)$, are determined using an adaptive scheme, see §V. For Gaussian white noise processes with $K = 2NW$ the expected squared weight is $\lambda_k/2NW$.

The multitaper usage of approximately bandlimited windows limits leakage from out-of-band frequencies, reducing bias, while the orthogonality of the windows guarantees a consistent estimate when $W$ is fixed.

**Distribution of Multitaper Estimators**

Multitaper estimators have $\chi^2_{2\alpha}$ distributions, where the effective degrees of freedom, $2\alpha$, can be approximated in terms of the weights

$$\alpha \approx \sum_{k=0}^{K-1} d_k^2(f) \approx K. \quad (A.8)$$

A particularly convenient standardized form of the spectrum

$$z(f) = \frac{\hat{S}(f)}{S(f)} \quad (A.9)$$

is distributed as a scaled central $\chi^2_{2\alpha}$ random variable, and takes a Gamma$(\alpha, 1/\alpha)$ probability distribution,

$$p(z; \alpha) = \frac{\alpha^\alpha}{\Gamma(\alpha)} z^{\alpha - 1} \exp(-\alpha z). \quad (A.10)$$

This distribution has mean 1 because $\mathbf{E}\{\hat{S}(f)\} = S(f)$ when the spectrum is constant and the variance is $1/\alpha$. The standardized spectrum will be used throughout this thesis. Standardized spectra are plotted on a linear scale instead of a log scale in the
A.3. STANDARDIZED SPECTRUM

y-axis.

Note that when the data contains a line component at frequency $f_0$, the distribution of the spectrum at $f_0$ changes from the null assumption of a central $\chi^2_{2\alpha}$, to a noncentral $\chi^2_{2\alpha}$ distribution, $[63,111]$. In this case $E\{\hat{S}(f_0)\} = \lambda + 2\alpha$, where $\lambda > 0$ is the noncentrality parameter.

A.3 A Note on the Standardized Spectrum

The denominator of $[A.9]$ is, in practice, estimated during the prewhitening step; it is helpful to think of it as the slowly-varying part of the spectrum, and $z(f)$ as the residual spectrum, $[118]$. The spectrum of a unit variance white noise process is identically 1 for $f \in (-1/2,1/2]$. The reason for standardizing the spectrum estimate using a Gamma distribution, as above, makes the residual after prewhitening of any process comparable to that of unit variance white noise. If prewhitening the spectrum has removed all of the structure in the process, then the standardized spectrum estimate should behave just as a Gamma($\alpha,1/\alpha$) random variable.

However, spectra containing a large number of line components, say $100\epsilon\%$ of the bins in the estimate, follows a mixture distribution of $100(1-\epsilon)\%$-central and $100\epsilon\%$-noncentral components. Identification of significant anomalies should be done using the central part of the distribution only, but because of the noncentral component, the prewhitening step will scale the distribution incorrectly (the base will be estimated too large). To correct for this, apply a robust $M-$estimate of scale for $\epsilon$-contaminated positive random variables as a correction factor, see $[70]$. Because the 5% point of the spectrum is likely to come entirely from samples in the central component, multiply the prewhitened estimate by the factor $Q(0.05;\alpha)/(q_{0.05} \ast 2\alpha)$, where $Q(0.05;\alpha)$ is
the 5% point of the Gamma($\alpha, 1/\alpha$) distribution, $q_{0.05}$ is the 5% quantile of the estimate, and $2\alpha$ scales the $\chi^2_{2\alpha}$ random variable to a Gamma($\alpha, 1/\alpha$) one. Unless the fraction of frequency bins in lines in the detail spectrum is larger than 95%, i.e. $\epsilon > 0.95$, significance in the standardized spectrum can safely be assessed according to a Gamma($K, 1/K$) distribution.

This discussion implies that careful prewhitening has been applied, see e.g. appendix D of [118] or [84]. When a power law is present in the spectrum, prewhitening consists of the removal of a linear trend in the log spectrum.

**A.4 A Note on Rayleigh Resolution**

The Rayleigh resolution, $R$, of a time series is the original lower bound on the spacing of peaks in the spectrum which can be resolved given the length of time over which the data sequence is observed. Hence, $R = 1/T$ Hz where $T$ is the length of the series (in seconds) or $R = 1/N\Delta t$ Hz when the series is sampled at intervals of length $\Delta t$ seconds (we assume unit sampling throughout). Note that the Rayleigh resolution does not depend on the sampling rate of the series, hence $R$ does not change with zero-padding or downsampling. In this paper, frequency is standardized in units of the Rayleigh resolution. Standardizing both the spectrum and the units of frequency is useful in order to construct standard reference tables for various derived quantities, which is done in Chapter 4 and Appendix B.

**A.5 Additional Tools**

The following additional, related, frequency-domain methods are used without introduction in the body of this thesis. More information about the harmonic F-test can
be found in, e.g. [111], and magnitude squared coherence can be read about in [106].

A.5.1 Harmonic F-test

The purpose of the F-test is to detect harmonic components, and it relies on the extended Munk-Hasselman representation for the process. Denoting \( dZ(f) \) as the usual stationary, orthogonal-increments process used in the Cramér spectral representation theorem, \( (A.4) \) for the stationary process \( x_t \), and extending the representation to allow for deterministic line components, write its first moment \( \mu(f) \) as

\[
\mu(f) = E\{dZ(f)\} = \sum_j \mu_j \delta(f - f_j)df
\]  

(A.11)

where \( \delta(f) \) is the Dirac delta function, and \( f_j \) are line component frequencies. The continuous part of the spectrum is \( S(f) df = E\{|dZ(f) - E\{dZ(f)\}|^2\} \), which is simply \( E\{|dZ(f)|^2\} \) when no harmonic components are present.

Estimation of \( \mu(f) \) can be done by minimizing the sum of squares distance between the eigencoeficients and their expected value. This minimization results in the following estimate for \( \mu \)

\[
\hat{\mu}(f) = \frac{\sum_{k=0}^{K-1} U_k(N,W;0)y_k(f)}{\sum_{k=0}^{K-1} U_k^2(N,W;0)}.
\]  

(A.12)

The variance-ratio gives an \( F \)-distributed test statistic as

\[
F(f) = \frac{(K - 1)|\hat{\mu}(f)|^2 \sum_{k=0}^{K-1} U_k(N,W;0)^2}{\sum_{k=0}^{K-1} |y_k(f) - \mu(f)U_k(0)|^2}
\]  

(A.13)

the numerator and denominator are \( \chi^2 \) distributed and have \( 2K - 2 \) degrees of
freedom (df), respectively when the line frequency is known. \( U_k(0) \) can be written in terms of Slepian sequences, see equation (A.2),

\[
U_k(N, W; 0) = \sum_{n=0}^{N-1} v_n^{(k)}(N, W).
\]

(A.14)

### A.5.2 Magnitude Squared Coherence

Coherency, \( \gamma(f) \) describes the covariance of two time series as a function of frequency in terms of magnitude and phase. The magnitude squared coherence, or MSC, is a partition of the covariance of the two signals into discrete frequency bins, similar to the way the spectrum partitions the variance of a single series. The multitaper estimator for coherence, \( c \), where \( x_t \) and \( y_t, \ t = 0, \ldots, N - 1 \) are two zero mean stationary time series with the same sampling rate and multitaper spectra \( \hat{S}_{xx}(f) \) and \( \hat{S}_{yy}(f) \) (most simply written without adaptive weighting), respectively, is given by

\[
c = \frac{\hat{S}_{xy}(f)}{\left[\hat{S}_{xx}(f)\hat{S}_{yy}(f)\right]^{1/2}} \quad (A.15)
\]

The numerator of this expression is the cross spectrum, which can be estimated by

\[
\hat{S}_{xy}(f) = \frac{1}{K} \sum_{k=0}^{K-1} \hat{x}_k(f)\hat{y}^*_k(f)
\]

(A.16)

where \( \hat{x}_k(f) \) and \( \hat{y}_k(f) \) are the eigencoefficients of the two series. When there exists a linear transformation with transfer function \( H(z) \) that relates the two series \( x(t) \) and \( y(t) \), the theoretical coherence is identically 1 over the entire frequency range.

The distribution of \( c \) is dependent upon the true value of the coherence \( \gamma^2 \) and is
given in [116, Eqn (2.53)]^2, or [49, p. 259]. In the special case where \( \gamma^2 = 0 \), there is no coherency between the two series, the probability distribution of \( c^2 \) is given by

\[
p(c^2) = (m - 1)(1 - c^2)^{(m-2)}.
\]  

(A.17)

A.5.3 Complex Demodulation

The process of complex demodulation allows one to isolate a small band in frequency and represent this bandlimited portion of the signal in the time domain. It involves first frequency-shifting a process by an amount \( f_c \), (that is, multiplying the process by the quantity \( e^{i2\pi f_c t} \) and then lowpass filtering to retain the frequency components within a certain (small) bandwidth \( W \) near the origin. The desired result is a bandlimited process, so it is sufficient to sample the series at steps \( \Delta = 1/2W \) apart in time.

One can accomplish complex demodulation in a number of ways, but the approach I have used throughout uses an expansion in Slepians described in [119, §V.D (39)], [114, p 17.], or [113]. Let \( \hat{x}(t; f) \) denote the process where the frequency \( f \) has been shifted to the origin, and expanded in Slepian sequences,

\[
\hat{x}(t; f) = \sum_{k=0}^{K-1} y_k(f) \sqrt{\lambda_k} v_i^{(k)}.
\]

(A.18)

and \( y_k(f) \) are the complex eigencoefficients of the process. Letting \( t_j = (j-1/2)/2W \), where \( j = 0, \ldots, 2NW \) one can downsample the result

\[
x \downarrow (j; f) = \hat{x}(t_j; f)
\]

(A.19)

---

Equation (2.53) has a minor typo, the factor \( (m - 1) \) should appear in the numerator.
to get $K$ samples of the new process. It may be helpful to think of as an expansion in Slepian sequences of the $K$ eigencoefficients of the process centered at a desired frequency $f$. Because the Slepian sequences form a basis for bandlimited processes, this expansion is optimally concentrated on the band $(f - W, f + W)$. The result is a complex process in time, and one generally plots the phase and the magnitude separately.

For more information on complex demodulation, consult [106, §6.2.1], or [13].
Appendix B

Standard Tables of Upcrossing Rates

Table B.1 shows the values of the constant $\psi$ given in equation (4.6) for various $K$, $NW$, and table B.2 on the next page gives the expected number of upcrossings of a multitaper estimate (4.5) for a given level per $10^5$ Rayleigh resolutions.

<table>
<thead>
<tr>
<th>$K$</th>
<th>3</th>
<th>3.5</th>
<th>4</th>
<th>4.5</th>
<th>5</th>
<th>5.5</th>
<th>6</th>
<th>6.5</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$C_R = NW$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>1.0248</td>
<td>0.9484</td>
<td>0.8869</td>
<td>0.8360</td>
<td>0.7930</td>
<td>0.7561</td>
<td>0.7238</td>
<td>0.6954</td>
<td>0.6701</td>
</tr>
<tr>
<td>2</td>
<td>1.0278</td>
<td>0.9502</td>
<td>0.8881</td>
<td>0.8369</td>
<td>0.7936</td>
<td>0.7565</td>
<td>0.7242</td>
<td>0.6957</td>
<td>0.6703</td>
</tr>
<tr>
<td>3</td>
<td>1.0346</td>
<td>0.9540</td>
<td>0.8904</td>
<td>0.8385</td>
<td>0.7948</td>
<td>0.7574</td>
<td>0.7249</td>
<td>0.6962</td>
<td>0.6707</td>
</tr>
<tr>
<td>4</td>
<td>1.0460</td>
<td>0.9612</td>
<td>0.8947</td>
<td>0.8412</td>
<td>0.7967</td>
<td>0.7588</td>
<td>0.7259</td>
<td>0.6971</td>
<td>0.6714</td>
</tr>
<tr>
<td>5</td>
<td>1.0284</td>
<td>0.9717</td>
<td>0.9022</td>
<td>0.8458</td>
<td>0.7997</td>
<td>0.7609</td>
<td>0.7275</td>
<td>0.6983</td>
<td>0.6724</td>
</tr>
<tr>
<td>6</td>
<td>0.8377</td>
<td>0.9514</td>
<td>0.9118</td>
<td>0.8533</td>
<td>0.8045</td>
<td>0.7641</td>
<td>0.7298</td>
<td>0.7000</td>
<td>0.6737</td>
</tr>
<tr>
<td>7</td>
<td>-</td>
<td>0.7740</td>
<td>0.8895</td>
<td>0.8620</td>
<td>0.8120</td>
<td>0.7690</td>
<td>0.7331</td>
<td>0.7024</td>
<td>0.6756</td>
</tr>
<tr>
<td>8</td>
<td>-</td>
<td>-</td>
<td>0.7230</td>
<td>0.8383</td>
<td>0.8199</td>
<td>0.7764</td>
<td>0.7381</td>
<td>0.7058</td>
<td>0.6781</td>
</tr>
<tr>
<td>9</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.6811</td>
<td>0.7951</td>
<td>0.7835</td>
<td>0.7453</td>
<td>0.7108</td>
<td>0.6815</td>
</tr>
<tr>
<td>10</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.6458</td>
<td>0.7580</td>
<td>0.7518</td>
<td>0.7179</td>
<td>0.6865</td>
</tr>
<tr>
<td>11</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.6155</td>
<td>0.7257</td>
<td>0.7237</td>
<td>0.6934</td>
</tr>
<tr>
<td>12</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.5892</td>
<td>0.6972</td>
<td>0.6986</td>
</tr>
<tr>
<td>13</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.5660</td>
<td>0.6718</td>
</tr>
<tr>
<td>14</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0.5455</td>
</tr>
</tbody>
</table>

Table B.1: Values for the constant $\psi$, or theoretical antecorrelation, for various choices of $NW, K$, calculated using $N = 10^4$. See also the supplement to [117].
<table>
<thead>
<tr>
<th>$C_R = 3.5$</th>
<th>$\alpha = 4$</th>
<th>$\alpha = 5$</th>
<th>$\alpha = 6$</th>
<th>$\alpha = 7$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P, %$</td>
<td>$z$</td>
<td>$U(z; 4)$</td>
<td>$\bar{w}(z)$</td>
<td>$z$</td>
</tr>
<tr>
<td>50</td>
<td>0.918</td>
<td>15417.2</td>
<td>3.243</td>
<td>0.934</td>
</tr>
<tr>
<td>90</td>
<td>1.670</td>
<td>6180.8</td>
<td>1.618</td>
<td>1.599</td>
</tr>
<tr>
<td>99</td>
<td>2.511</td>
<td>891.0</td>
<td>1.122</td>
<td>2.321</td>
</tr>
<tr>
<td>99.9</td>
<td>3.266</td>
<td>109.3</td>
<td>0.915</td>
<td>2.959</td>
</tr>
<tr>
<td>99.99</td>
<td>4.666</td>
<td>1.4</td>
<td>0.712</td>
<td>4.130</td>
</tr>
<tr>
<td>99.999</td>
<td>5.338</td>
<td>0.2</td>
<td>0.651</td>
<td>4.686</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$C_R = 4.0$</th>
<th>$\alpha = 5$</th>
<th>$\alpha = 6$</th>
<th>$\alpha = 7$</th>
<th>$\alpha = 8$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P, %$</td>
<td>$z$</td>
<td>$U(z; 5)$</td>
<td>$\bar{w}(z)$</td>
<td>$z$</td>
</tr>
<tr>
<td>50</td>
<td>0.934</td>
<td>1445.8</td>
<td>3.461</td>
<td>0.945</td>
</tr>
<tr>
<td>90</td>
<td>1.599</td>
<td>5844.6</td>
<td>1.711</td>
<td>1.546</td>
</tr>
<tr>
<td>99</td>
<td>2.321</td>
<td>845.3</td>
<td>1.183</td>
<td>2.185</td>
</tr>
<tr>
<td>99.9</td>
<td>2.959</td>
<td>103.8</td>
<td>0.963</td>
<td>2.742</td>
</tr>
<tr>
<td>99.99</td>
<td>3.556</td>
<td>12.0</td>
<td>0.835</td>
<td>3.261</td>
</tr>
<tr>
<td>99.999</td>
<td>4.130</td>
<td>1.3</td>
<td>0.749</td>
<td>3.756</td>
</tr>
<tr>
<td>99.9999</td>
<td>4.686</td>
<td>0.1</td>
<td>0.686</td>
<td>4.235</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$C_R = 5.0$</th>
<th>$\alpha = 7$</th>
<th>$\alpha = 8$</th>
<th>$\alpha = 9$</th>
<th>$\alpha = 10$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P, %$</td>
<td>$z$</td>
<td>$U(z; 7)$</td>
<td>$\bar{w}(z)$</td>
<td>$z$</td>
</tr>
<tr>
<td>50</td>
<td>0.953</td>
<td>12977.4</td>
<td>3.853</td>
<td>0.959</td>
</tr>
<tr>
<td>90</td>
<td>1.505</td>
<td>5314.1</td>
<td>1.882</td>
<td>1.471</td>
</tr>
<tr>
<td>99</td>
<td>2.082</td>
<td>772.3</td>
<td>1.295</td>
<td>2.000</td>
</tr>
<tr>
<td>99.9</td>
<td>2.580</td>
<td>95.1</td>
<td>1.052</td>
<td>2.453</td>
</tr>
<tr>
<td>99.99</td>
<td>3.041</td>
<td>11.0</td>
<td>0.912</td>
<td>2.870</td>
</tr>
<tr>
<td>99.999</td>
<td>3.480</td>
<td>1.2</td>
<td>0.817</td>
<td>3.265</td>
</tr>
<tr>
<td>99.9999</td>
<td>3.903</td>
<td>0.1</td>
<td>0.748</td>
<td>3.645</td>
</tr>
</tbody>
</table>

Table B.2: Expected number of upcrossings $U(z, K)$ and dwell bandwidths $\bar{w}(z)$ for three common choices of $NW$. For more choices of $NW, K$, consult the supplementary material to \[117\].
<table>
<thead>
<tr>
<th>$C_R = 6$</th>
<th>$\alpha = 9$</th>
<th>$\alpha = 10$</th>
<th>$\alpha = 11$</th>
<th>$\alpha = 12$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P,%$</td>
<td>$z$</td>
<td>$U(z; 9)$</td>
<td>$\bar{w}(z)$</td>
<td>$z$</td>
</tr>
<tr>
<td>50</td>
<td>0.963</td>
<td>11900.4</td>
<td>4.202</td>
<td>0.967</td>
</tr>
<tr>
<td>90</td>
<td>1.444</td>
<td>4911.4</td>
<td>2.036</td>
<td>1.421</td>
</tr>
<tr>
<td>99</td>
<td>1.934</td>
<td>716.3</td>
<td>1.396</td>
<td>1.878</td>
</tr>
<tr>
<td>99.9</td>
<td>2.351</td>
<td>88.3</td>
<td>1.133</td>
<td>2.266</td>
</tr>
<tr>
<td>99.99</td>
<td>3.093</td>
<td>0.1</td>
<td>0.879</td>
<td>2.952</td>
</tr>
<tr>
<td>99.999</td>
<td>3.440</td>
<td>0.1</td>
<td>0.804</td>
<td>3.271</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$C_R = 8$</th>
<th>$\alpha = 13$</th>
<th>$\alpha = 14$</th>
<th>$\alpha = 15$</th>
<th>$\alpha = 16$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P,%$</td>
<td>$z$</td>
<td>$U(z; 13)$</td>
<td>$\bar{w}(z)$</td>
<td>$z$</td>
</tr>
<tr>
<td>50</td>
<td>0.974</td>
<td>10390.9</td>
<td>4.812</td>
<td>0.976</td>
</tr>
<tr>
<td>90</td>
<td>1.368</td>
<td>4331.5</td>
<td>2.309</td>
<td>1.354</td>
</tr>
<tr>
<td>99</td>
<td>1.755</td>
<td>634.8</td>
<td>1.575</td>
<td>1.724</td>
</tr>
<tr>
<td>99.9</td>
<td>2.079</td>
<td>78.4</td>
<td>1.275</td>
<td>2.032</td>
</tr>
<tr>
<td>99.99</td>
<td>2.371</td>
<td>9.1</td>
<td>1.102</td>
<td>2.309</td>
</tr>
<tr>
<td>99.999</td>
<td>2.645</td>
<td>1.0</td>
<td>0.987</td>
<td>2.568</td>
</tr>
<tr>
<td>99.9999</td>
<td>2.906</td>
<td>0.1</td>
<td>0.903</td>
<td>2.815</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$C_R = 10$</th>
<th>$\alpha = 17$</th>
<th>$\alpha = 18$</th>
<th>$\alpha = 19$</th>
<th>$\alpha = 20$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$P,%$</td>
<td>$z$</td>
<td>$U(z; 17)$</td>
<td>$\bar{w}(z)$</td>
<td>$z$</td>
</tr>
<tr>
<td>50</td>
<td>0.980</td>
<td>9358.4</td>
<td>5.343</td>
<td>0.982</td>
</tr>
<tr>
<td>90</td>
<td>1.321</td>
<td>3925.7</td>
<td>2.547</td>
<td>1.311</td>
</tr>
<tr>
<td>99</td>
<td>1.649</td>
<td>577.3</td>
<td>1.732</td>
<td>1.628</td>
</tr>
<tr>
<td>99.9</td>
<td>1.919</td>
<td>71.4</td>
<td>1.400</td>
<td>1.888</td>
</tr>
<tr>
<td>99.99</td>
<td>2.161</td>
<td>8.3</td>
<td>1.209</td>
<td>2.121</td>
</tr>
<tr>
<td>99.999</td>
<td>2.386</td>
<td>0.9</td>
<td>1.082</td>
<td>2.337</td>
</tr>
<tr>
<td>99.9999</td>
<td>2.600</td>
<td>0.1</td>
<td>0.989</td>
<td>2.542</td>
</tr>
</tbody>
</table>

Table B.3: (Continued) Expected number of upcrossings $U(z, K)$ and dwell bandwidths $\bar{w}(z)$ for three common choices of $C_R = NW$. For more choices of $NW, K$, consult the supplementary material to [117].
Appendix C

Karhunen Loève Expansion

The Karhunen-Loève expansion is a useful tool in digital signal processing applications for decomposing time series into a linear combination of orthogonal series with uncorrelated coefficients. It is closely related to principal components analysis and the singular value decomposition. The approach and notation presented here is from [111]; an accessible introduction is available in [122, Chapter 6], also [53, 72].

The Karhunen Loève expansion is defined for \( N \) samples of a time series with autocovariance function \( R(\tau) \) as

\[
X = \Psi C; \quad C = \Psi^H X \tag{C.1}
\]

where \( C \) is the coefficient vector and \( \Psi \) is the matrix whose \( m \)th column is the eigenvector associated with the eigenvalue \( \theta_n \) in the matrix equation

\[
\Psi \Theta = R \Psi \tag{C.2}
\]

where \( R \) is the Toeplitz matrix of autocovariances (Toeplitz matrices obey the relation
\( r_{jk} = r_{j+1,k+1} \) for all \( j, k = 0, \ldots, N - 2 \), or equivalently

\[
\theta_n \psi_n(t) = \sum_{s=0}^{n-1} R(t - s) \psi_n(s) \tag{C.3}
\]

and the \( \psi \)'s are normalized such that \( \sum_t \psi_n(t) \psi_k(t) = \delta_{n,k} \). The expansion coefficients \( c_n \)

\[
c_n = \sum_{t=0}^{N-1} \psi_n(t)x(t) \tag{C.4}
\]

have expected value 0 and covariance \( \mathbb{E}\{c_n c_m\} = \theta_n \delta_{n,m} \).
Appendix D

Theorems Related to Eigenvalue Problems

In this brief appendix, I will give two theorems central to the study of Fredholm-type integral equations, Mercer’s theorem and Szegö’s theorem, and their discrete analogs.

D.1 Mercer’s Theorem

A symmetric kernel $R(s, t)$ positive, continuous, and such that the integral
\[ \int_{-\infty}^{\infty} |R(s, t)|^2 ds \] is bounded on $(-\infty, \infty)$, is the sum of its Fourier series which is absolutely and uniformly convergent with respect to the pair of variables $s, t$

\[ R(s, t) = \sum_{n=1}^{\infty} \frac{\psi_n(s)\psi_n(t)}{\theta_n}. \] (D.1)

For a review of integral equations and their solutions, consult e.g. [85].

D.2 Szegö’s Theorem

Szegö’s theorem is an extremely useful result concerning eigenvalues of Toeplitz matrices, see [108, 18, 45] but especially Grenander and Szegö [46, §5.2]. An important
corollary of Szegö’s theorem is that there is an asymptotic relation between the eigenvalues of the autocovariance matrix of a stationary process and the sorted spectrum, where the spectrum is sampled between 0 to 1/2.

Let \( x(t) \) be a real, normally distributed, stationary process with spectral density \( S(f) \), as we have throughout. The Weiner-Khintchine theorem gives the relation

\[
R(t - u) = \int_{-1/2}^{1/2} S(f)e^{i(t-u)f} df, \quad \text{for } t, u = 0, \ldots, N - 1.
\] (D.2)

between the autocovariance and the spectrum. Define the \( N \)-sample Toeplitz autocovariance matrix \( R_N = [R(t - u)]_{t, u=0}^{N-1} \). The matrix \( R_N \) is positive-definite and has real, positive eigenvalues for all \( N \). Let the eigenvalues of \( R_N \) be denoted \( \{\lambda^{(N)}_n\}_{n=0}^{N-1} \) and let them be ordered \( \lambda^{(N)}_0 \geq \lambda^{(N)}_1 \geq \ldots \geq \lambda^{(N)}_{N-1} > 0 \) for convenience.

The theorem of Szegö is as follows: if \( S(f) \in L^1 \) is bounded above and below by \( m \) and \( M \), respectively, for \( f \in [-1/2, 1/2) \), and \( F(\cdot) \) is a continuous function defined on \((m, M)\), the following asymptotic formula holds

\[
\lim_{N \to \infty} \frac{F(\lambda^{(N)}_0) + F(\lambda^{(N)}_1) + \ldots + F(\lambda^{(N)}_{N-1})}{N + 1} = \int_{-1/2}^{1/2} F(S(f)) df.
\] (D.3)

and \( M \geq \lambda^{(N)}_0 \geq \lambda^{(N)}_1 \geq \ldots \geq \lambda^{(N)}_{N-1} \geq m > 0 \). Stationary processes with Lorentzian spectra have the peculiar property that the eigenvalues \( \theta_n \) are actually equal to the spectrum at the frequencies \( b_n - \omega_0 \), see [122] p 543 remark 1 and §6.4.7. With some additional arguments, see [122] §6.4.7, it is true in general that the eigenvalues of the covariance matrix of a stationary process approach the spectrum, in a certain sense.