

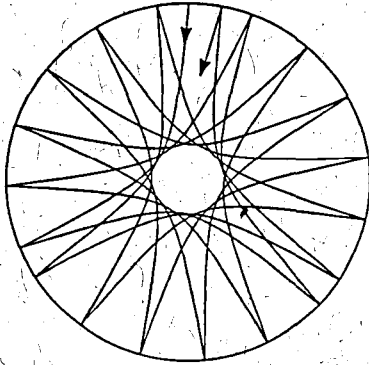
NASA-CR-175661

JPL PUBLICATION 84-84

NASA-CR-175661  
19850016749

# Solar Seismology From Space

## A Conference at Snowmass, Colorado



December 15, 1984

**NASA**

National Aeronautics and  
Space Administration

**Jet Propulsion Laboratory**  
California Institute of Technology  
Pasadena, California

**LIBRARY COPY**

JUL 17 1985

LANGLEY RESEARCH CENTER  
LIBRARY, NASA  
HAMPTON, VIRGINIA



NF00477



DEPARTMENT OF ASTRONOMY  
LOS ANGELES, CALIFORNIA 90024

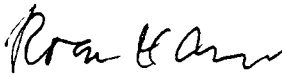
February 13, 1985

Dear Colleague:

Enclosed is a copy of the proceedings of a conference held at Snowmass, Colorado, September, 1983. This conference was part of a study sponsored by NASA to define a possible space experiment to measure solar oscillations from space. A separate working group report has recently been released by NASA describing such an experiment and the advantages of basing this experiment in space. I am sending this copy of the conference proceedings to you in your capacity as a representative of your institution so that this volume can be included in your library's archive collection. If you have not already received a personal copy of this volume and would wish to have one, please write to Tom Thorpe at JPL, Pasadena, CA 91109, or to me at UCLA, Los Angeles, CA 90024. It is my hope that the papers in this volume will be readily available both to those we know are interested in the subject as well as those people who may become interested in the future.

Thank you for your help.

Sincerely yours,

  
Roger K. Ulrich

R.K.U./rlo

JPL PUBLICATION 84-84

# **Solar Seismology From Space**

## **A Conference at Snowmass, Colorado**

**R.K. Ulrich, Chairman**  
**J. Harvey**  
**E.J. Rhodes, Jr.**  
**J. Toomre**  
Science Organizing Committee

December 15, 1984



National Aeronautics and  
Space Administration

**Jet Propulsion Laboratory**  
California Institute of Technology  
Pasadena, California

1175-25060 #

This publication was prepared by the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology.



## ABSTRACT

The scientific conference at Snowmass, Colorado, on solar seismology was the result of a recommendation by the NASA sponsored Science Working Group on the Study of Solar Oscillations from Space. Solar or Helio Seismology is a subject of remarkable maturity given its relative youth. Several exciting discoveries have been made with ground-based techniques, and reports of some of these are included in these proceedings. The quality of the ground-based observing environment suffers from several degrading factors: diurnal interruptions and thermal variations, atmospheric seeing and transparency fluctuations and adverse weather interruptions are among the chief difficulties. The limited fraction of the solar surface observable from only one vantage point is also a potential limitation to the quality of the data available without going to space. The significance of all these factors is not immediately apparent, and this conference was intended to provide a forum permitting a free interchange of ideas and analysis. Primary conference goals were to discuss in depth the scientific return from current observations and analyses of solar oscillations, to discuss the instrumental and site requirements for realizing the full potential of the seismic analysis method, and to help bring new workers into the field by collecting and summarizing the key background theory. At the conclusion of the conference there was a clear consensus that ground-based observation would not be able to provide data of the quality required to permit a substantial analysis of the solar convection zone dynamics or to permit a full deduction of the solar interior structure.



**SOLAR SEISMOLOGY FROM SPACE**

**A CONFERENCE AT SNOWMASS COLORADO**

**Sponsored by NASA and JPL**

**August 17 to 19, 1983**

**TABLE OF CONTENTS**

**PART 1**

**INTRODUCTION - METHODS AND GOALS**

- |    |  |    |
|----|--|----|
| 1. | CONFERENCE SUMMARY<br>Roger K. Ulrich  | 3  |
| 2. | OVERVIEW OF SOLAR SEISMOLOGY: OSCILLATIONS AS PROBES OF<br>INTERNAL STRUCTURE AND DYNAMICS IN THE SUN<br>Juri Toomre                                   | 7  |
| 3. | WHAT WOULD A DYNAMO THEORIST LIKE TO KNOW ABOUT THE<br>DYNAMICS OF THE SOLAR CONVECTION ZONE?<br>Peter A. Gilman                                       | 41 |
| 4. | SOLAR INVERSE THEORY<br>Douglas Gough  | 49 |
| 5. | ROTATIONAL INVERSION FROM GLOBAL SOLAR OSCILLATIONS<br>Jorgen Christensen-Dalsgaard and Douglas Gough  | 79 |
| 6. | SENSITIVITY OF INFERRED SUBPHOTOSPHERIC VELOCITY FIELD TO<br>MODE SELECTION, ANALYSIS TECHNIQUE AND NOISE<br>Frank Hill, Douglas Gough and Juri Toomre | 95 |

**PART 2**

**INSTRUMENTS AND OBSERVATIONS**

- |    |  |     |
|----|--|-----|
| 7. | THE MAGNETO-OPTICAL FILTER, WORKING PRINCIPLES AND<br>RECENT PROGRESS<br>A. Cacciani and E. J. Rhodes, Jr. | 115 |
|----|--|-----|

8.	EVALUATION OF A MAGNETO-OPTICAL FILTER AND A FABRY-PEROT INTERFEROMETER FOR THE MEASUREMENT OF SOLAR VELOCITY FIELDS FROM SPACE	125
	Edward J. Rhodes, Jr., Alessandro Cacciani, Jacques Blamont, Steven Tomczyk, Roger K. Ulrich and Robert F. Howard	
9.	THE FOURIER TACHOMETER II -- AN INSTRUMENT FOR MEASURING GLOBAL SOLAR VELOCITY FIELDS	157
	Timothy M. Brown	
10.	OBSERVATIONS OF INTERMEDIATE-DEGREE SOLAR OSCILLATIONS	165
	J. W. Harvey and T. L. Duvall, Jr.	
11.	DETECTION OF SOLAR GRAVITY MODE OSCILLATIONS	173
	Philip H. Scherrer	
12.	SOLAR GRAVITY MODES FROM ACRIM/SMM IRRADIANCE DATA	183
	Claus Frohlich and Philippe Delache	
13.	OBSERVATIONS OF LOW-DEGREE MODES FROM THE SOLAR MAXIMUM MISSION	195
	Martin Woodard	
14.	IMPLICATIONS OF OBSERVED FREQUENCIES OF SOLAR p MODES	199
	Jorgen Christensen-Dalsgaard and Douglas Gough	

**PART 3  
ANALYSIS AND INTERPRETATION OF OBSERVATIONS,  
SPATIAL FILTERS, SEEING AND DATA GAP FILLING**

15.	OBSERVED SPATIAL PROPERTIES OF THE SOLAR EIGENFUNCTIONS AND THE IMPLICATIONS FOR THE EXISTENCE OF RESOLVED MULTIPLETS	207
	Thomas P. Caudell, Henry A. Hill and Randall J. Bos	
16.	OPTIMIZED RESPONSE FUNCTIONS FOR TWO-DIMENSIONAL OBSERVATIONS OF SOLAR OSCILLATIONS	219
	Jorgen Christensen-Dalsgaard	
17.	THE EFFECTS OF IMAGE MOTION ON THE $\rho - \nu$ DIAGRAM	255
	Frank Hill	
18.	THE EFFECTS OF SEEING ON NOISE	263
	Roger K. Ulrich, Edward J. Rhodes, Jr., Alessandro Cacciani and Steven Tomczyk	
19.	THE EFFECTS OF A NEARLY 100% DUTY CYCLE ON OBSERVATIONS OF SOLAR OSCILLATIONS	271
	Frank Hill	

20.	ANALYSIS AND INTERPRETATION OF SYNTHETIC TIME STRINGS OF OSCILLATION DATA	279
	Barbara W. Mihalas, Jorgen Christensen-Dalsgaard and Timothy M. Brown	
21.	SPECTRAL INFORMATION FROM GAPPED DATA: A COMPARISON OF TECHNIQUES	293
	J. R. Kuhn	
22.	METHODS FOR CLOUD COVER ESTIMATION	305
	David L. Glackin, James R. Huning, Jeffrey H. Smith and Thomas L. Logan	

**PART 4  
EXCITATION, DAMPING AND ROTATIONAL SPLITTING  
OF THE OSCILLATIONS**

23.	NUMERICAL SIMULATIONS OF CONVECTIVELY EXCITED GRAVITY WAVES	315
	G. A. Glatzmaier	
24.	ON THE INFLUENCE OF TURBULENT MOTIONS ON NON-RADIAL OSCILLATIONS	325
	E. R. Durney	
25.	THE STABILITY OF THE LOW DEGREE FIVE MINUTE SOLAR OSCILLATIONS	335
	Russell B. Kidman and Arthur N. Cox	
26.	AN EXCITATION MECHANISM FOR SOLAR FIVE-MINUTE OSCILLATIONS OF INTERMEDIATE AND HIGH DEGREE	345
	H. M. Antia, S. M. Chitre and D. Narasimha	
27.	THE RESONANT COUNT DIAGRAM AND SOLAR $g$ MODE OSCILLATIONS	349
	D. B. Guenther and P. Demarque	
28.	FINE STRUCTURE OF SOLAR ACOUSTIC OSCILLATIONS DUE TO ROTATION	351
	Philip R. Goode and W. Dziembowski	

**PART 5**  
**THE SOLAR EQUATION OF STATE**

- |   |     |
|---|-----|
| 29. IONIZATION EQUILIBRIUM AND EQUATION OF STATE IN THE<br>SOLAR INTERIOR<br>Forrest J. Rogers  | 357 |
| 30. THE SENSITIVITY OF SOLAR EIGENFREQUENCIES TO THE TREATMENT<br>OF THE EQUATION OF STATE<br>Roger K. Ulrich and Edward J. Rhodes, Jr. | 371 |

PART 1

INTRODUCTION — METHODS AND GOALS





## CONFERENCE SUMMARY

Roger K. Ulrich  
Department of Astronomy,  
University of California, Los Angeles, CA 90024, U.S.A.

Solar or Helio Seismology is a subject of remarkable maturity given its relative youth. Several exciting discoveries have been made with ground-based techniques, and reports of some of these are included in these proceedings. The quality of the ground-based observing environment suffers from several degrading factors: diurnal interruptions and thermal variations, atmospheric seeing and transparency fluctuations and adverse weather interruptions are among the chief difficulties. The limited fraction of the solar surface observable from only one vantage point is also a potential limitation to the quality of the data available without going to space. The significance of all these factors is not immediately apparent and in order to provide a forum to permit a free interchange of ideas and analysis, the NASA sponsored Science Working Group on the Study of Solar Oscillations from Space recommended at its January 1983 meeting at JPL that a scientific conference on solar seismology be held. This conference at Snowmass was the result of that recommendation. I agreed to serve as the principal organizer of this conference and J. Harvey, J. Toomre and E. Rhodes agreed to assist me. The thoughtful and helpful advice that they provided was invaluable in the success of the conference.

Our twin goals were to discuss in depth the scientific return from current observations and analyses of solar oscillations and to discuss the instrumental and site requirements for realizing the full potential of the seismic analysis method. The organization of the conference was intended to highlight these two goals starting with some of the most challenging and exciting scientific objectives. We also felt it important to help bring new workers into the field by collecting and summarizing the key background theory. The introductory chapter by Toomre provides such an overview and tutorial for those readers not familiar with the subject. Toomre describes the nature of the oscillations and reviews the status of the principal observational results. The following paper by Gilman is meant primarily as a challenge to the observers - existing observations are not close to providing the quality of data necessary to answer Gilman's questions about the solar dynamo; however, we hope that such dynamo theory questions probably could be answered with better data. The technique of deducing the internal dynamics of the solar envelope from the oscillations involves the inversion of an integral equation. The series of papers by Gough, Christensen-Dalsgaard, Hill and Toomre (in various combinations and permutations) sets forth the theory needed to carry out this inversion. These authors show that the inversion process is vulnerable to noise in the data because of the need to take differences between nearly identical quantities. Given adequate data, they also show that a remarkable amount of detailed information about solar dynamics can be deduced.

The measurement of solar oscillations requires a stable instrument which ideally permits the determination of the velocity pattern over the solar surface within a short period of time. Most early work which measured solar velocities utilized spectroscopic techniques with a scanning slit. With such a system an image was gradually built by rastering over the solar surface. The time delay and inefficiency of the older method as well as the size and weight of high resolution spectrographs has led to the development of newer methods of velocity measurement which in essence involve tunable narrow band filters and an array detector. Three such approaches are discussed by Rhodes, Caccaini et al and by Brown. Promising results are described in these papers. Applications of the more mature spectrographic technique are presented in the papers by Harvey and Duvall and by Scherrer. The quality of these results is good. The papers by Frohlich and Delache and by Woodard give the first results from a space-borne experiment. Their serendipitous detection of the solar oscillations has provided us an opportunity to study data which is unaffected by the presence of day-night gaps although the observing sequence is shorter: 55 minutes on and 50 minutes off cycle.

The next part of the conference addressed two deficiencies with ground-based observations -- the degrading effect of the earth's atmosphere and the difficulties caused by gaps resulting from the diurnal observing cycle. The discussions of the filter functions by Christensen-Dalsgaard and by Caudell et al are indirectly related to the atmosphere degradation since the filter functions must function through the atmospheric distortion. The theoretical analysis of the seeing problem by Hill, et al and the observational analysis by Ulrich, et al both indicate that the atmospheric degradation reduces the sensitivity to the highest  $\lambda$  values and introduces a background noise for all values of  $\lambda$ . This latter difficulty is particularly important because as is shown in the section on inverse theory, a variety of modes including those which have low excitation amplitudes will be needed to obtain sensitivity to the structure and dynamics throughout the sun. The papers on the filling of temporal gaps by Hill, by Mihalas, Christensen-Dalsgaard and Brown and by Kuhn show that as long as the duty cycle of the observations exceeds about 80%, nearly perfect power spectra can be recovered. This conclusion of course is sensitive to the level of noise in the data. One solution to the day-night problem is to take data from a global network of stations. The siting of such a network will require some consideration of cloudiness at prospective locations. The paper by Glackin et al gives some guidance on the utilization of space-based and other data to estimate the cloudiness.

The ultimate limit on the accuracy of the solar oscillation data will be set by the intrinsic lifetime of the modes of oscillation. The theory of mode excitation and mode-mode coupling which might some day be able to provide guidance on this topic is presently in a rather underdeveloped state. The papers by Guenther and Demarque, by Glatzmaier, by Durney, by Kidman and Cox and by Antia et al discuss various aspects of this excitation problem. Another area where some limits on the quality of the deduction might arise is through the solar equation of state. Rogers presents an introduction to the current state of the art in this area and the paper by Ulrich and Rhodes gives some preliminary estimation of the possible importance of the equation of state uncertainties.

Following the scheduled part of the meeting the conference participants held an informal discussion of the question:

'Why should oscillations be studied from space?'

The general consensus was that the atmospheric seeing problem appears to be the most fundamental limitation to the ground-based observations. Included in this problem are the effects of the image distortion and the effects of transparency variations. The uncertainties posed in the selection of global network sites represent a significant risk which can only be evaluated convincingly by building such a network. The readily available data on cloudiness and atmospheric seeing are inadequate to permit a highly reliable determination of site quality without gathering new observations. The questions of the dynamics of the solar convection zone and the relationship of these dynamics to the solar activity can only be answered with data having a high degree of continuity and contiguity. Further, it is essential to have regular access to oscillation modes possessing small horizontal wavelengths. There was a clear consensus that the ground-based network would not be able to provide data of the quality required to permit a substantial analysis of the solar convection zone dynamics. Such probing is essential to come to understand what controls the flow and magnetic structures in rich evidence at the solar surface. It is also clear that all acoustic oscillations are influenced by conditions at the upper reflection point and the stratification there must be resolved unambiguously by using modes with short wavelengths that are observable only from space. Finally, a space observatory affords the best way to measure those long period acoustic and gravity modes that sample the deep interior of the sun but which are of very modest amplitude in the atmosphere.



**OVERVIEW OF SOLAR SEISMOLOGY: OSCILLATIONS AS  
PROBES OF INTERNAL STRUCTURE AND DYNAMICS IN THE SUN**

**JURI TOOMRE**

Department of Astrophysical, Planetary and Atmospheric  
Sciences, and Joint Institute for Laboratory Astrophysics,  
University of Colorado, Boulder, CO 80309, U.S.A.

**ABSTRACT**

The physical nature of solar oscillations is reviewed. The nomenclature of the subject and the techniques used to interpret the oscillations are discussed. Many of the acoustic and gravity waves that can be observed in the atmosphere of the Sun are actually resonant or standing modes of the interior; precise measurements of the frequencies of such modes allow deductions of the internal structure and dynamics of this star. The scientific objectives of such studies of solar seismic disturbances, or of solar seismology, will be outlined. The reasons for why it would be very beneficial to carry out further observations of solar oscillations both from ground-based networks and from space will be discussed.

**1. INTRODUCTION**

Observations of the atmosphere of the Sun reveal a broad range of acoustic and gravity waves, with many of these representing internal modes of oscillation which penetrate from below into the atmosphere. They are detectable in the solar photosphere as Doppler shifts of spectral lines or as associated intensity fluctuations. These seismic disturbances of the Sun, much like those of the Earth, afford the possibility of studying various aspects of the internal structure and dynamics of this star. This has led to the rapidly evolving subject called solar seismology or helioseismology. We will here briefly review the notation and some of the physical reasoning that is essential for making sense of this subject. The reader may also find it convenient to turn to several current review articles on solar seismology (e.g. Leibacher and Stein 1981; Gough 1983; Deubner and Gough 1984; Brown, Mihalas and Rhodes 1984).

Two of the main restoring forces for solar oscillations are pressure and buoyancy. These yield two distinct types of wave modes: the ones commonly labelled p modes are dominated by pressure and consist of propagating acoustic waves, while those labelled g modes are dominated by buoyancy and consist of interval gravity waves. The

p modes have relatively short periods (typically about 5 minutes), while the g modes have much longer periods (typically several hours). In general a given mode of oscillation is trapped within a specific region of the Sun, in the sense that most of its energy is in that region. Accordingly, the temporal frequency  $\nu$  of the mode is largely determined by the stratification of that region. For p modes the trapping is between an upper turning point just beneath the photosphere where the mode is reflected by the rapid decrease in the density, and a lower turning point where the mode is turned around more gradually by the increase in sound speed with depth. The characteristic period of a resonant acoustic wave within such a cavity depends upon its travel time between the turning points, and this is controlled by the varying sound speed  $c$  at which that wave propagates. For g modes, wave propagation is generally possible only in regions below the convection zone (which occupies the outer 30% of the Sun by radius) where buoyancy can be a restoring force rather than a destabilizing one. A particular g mode is trapped in regions where its frequency  $\nu$  is less than the local buoyancy frequency  $N$ . Once again, the characteristic period of a gravity mode depends upon its travel time through the region of trapping, and this is controlled by the variation of  $N$  with depth. By determining the frequencies of such p and g modes with precision, one is able to probe either the variations of sound speed  $c$  or of buoyancy frequency  $N$  over different intervals within the Sun. Since  $c$  and  $N$  measure differing aspects of the stratification and the composition of the interior, solar seismology can serve to guide and validate theories for the internal structure and the evolution of a star like the Sun. Clearly the frequencies of these global modes of oscillation are integral properties of the stratification within the Sun. Given a sufficient variety of observable modes, much about the internal constitution of the Sun can be unravelled.

## 2. SCIENTIFIC GOALS OF SOLAR SEISMOLOGY

Seismic probing of the Sun affords the potential of measuring for the first time the static and dynamical structure of the interior of a star. As a consequence, solar seismology can be expected to make major contributions in the following areas of what might be called "big physics":

- a) Determination of the initial solar helium abundance. An important prediction of all cosmological models is the proportionate amount of helium produced at the formation of the universe. The distribution of helium within the current Sun can be inferred from the frequencies of the global modes of oscillation, thereby permitting detailed estimates of the abundance of helium in the protosolar state. This may yield the most reliable estimates of helium abundance in the early universe, thereby impacting both theories of cosmology and stellar evolution.
- b) Solar neutrinos. The low observed neutrino flux could be due to an error in solar models or in particle physics. Direct probing of the solar interior with p and g modes, particularly near its center, may be able to resolve this long-standing puzzle.

c) Stellar interior dynamics. The differential rotation of the Sun with both depth and latitude can be deduced from the fine structure of the oscillation frequencies. Flows associated with giant convection cells below the surface can be similarly probed. The possible presence of very strong magnetic fields of primordial origin deep within the Sun would also produce fine structure in the frequencies. Further, the depth of the convection zone and its stratification can be tested by precise determination of the basic frequencies of the modes. Thus one may begin to address questions of vital importance to theories for convection, stellar evolution, and stellar activity cycles.

d) Solar gravitational potential. Tests of general relativity involving the orbit of Mercury depend upon knowing the quadrupole component  $J_2$  of the Sun's gravitational field. By measuring the internal rotation of the Sun using the oscillation modes, one can deduce departures from spherical symmetry of its gravitational potential. This may resolve questions of the accuracy of the theory of general relativity.

Table 1 outlines the scientific objectives of solar seismology, also indicating just which measurable properties of the oscillation frequencies will be used to address those objectives. We turn next to a brief overview of some of the results to date from solar seismology that bear on those goals.

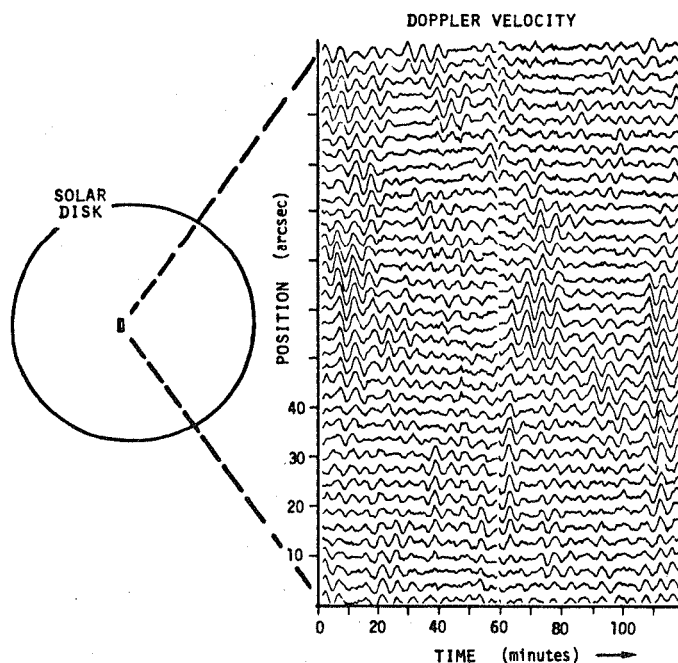
Table 1. SCIENTIFIC OBJECTIVES OF OSCILLATIONS AS PROBES OF INTERIOR STRUCTURE

<u>MEASURED QUANTITY</u>	<u>DERIVED SOLAR PROPERTIES</u>
<u>FREQUENCY:</u> $\nu_{n,\ell}$	DETERMINE STRATIFICATION OF DENSITY, TEMPERATURE AND PRESSURE WITH DEPTH <ul style="list-style-type: none"> <li>o CALIBRATE STELLAR STRUCTURE THEORY</li> <li>o OBTAIN DEPTH OF CONVECTION ZONE</li> </ul>
<u>P MODES:</u> SAMPLE SOUND SPEED C WITH DEPTH	DETERMINE CHEMICAL COMPOSITION AND ITS GRADIENTS <ul style="list-style-type: none"> <li>o ESTIMATE PRIMORDIAL HELIUM ABUNDANCE</li> <li>o TEST MIXING IN INTERIOR</li> <li>o ATTEMPT TO RESOLVE SOLAR NEUTRINO PROBLEM</li> </ul>
<u>G MODES:</u> SAMPLE BUOYANCY FREQUENCY N WITH DEPTH	
<u>FREQUENCY FINE STRUCTURE:</u> $\nu_{n,\ell,m}$	DETERMINE DIFFERENTIAL ROTATION: ANGULAR VELOCITY $\Omega$ WITH RADIUS AND LATITUDE <ul style="list-style-type: none"> <li>o OBTAIN SOLAR QUADRUPOLE COMPONENT <math>J_2</math></li> <li>o TEST SPINDOWN IN STARS AS THEY EVOLVE</li> </ul>
	DETERMINE COUPLING OF CONVECTION WITH ROTATION <ul style="list-style-type: none"> <li>o PROVIDE VITAL INPUT TO SOLAR DYNAMO MODELS</li> </ul>
	SEEK EVIDENCE FOR STRONG FOSSIL MAGNETIC FIELD
<u>VARIATION WITH TIME OF FREQUENCY FINE STRUCTURE:</u> $\nu_{n,\ell,m}$ (TIME)	PROBE VELOCITY AND TEMPERATURE STRUCTURES OF GIANT CELLS IN CONVECTION ZONE <ul style="list-style-type: none"> <li>o TEST DYNAMICS OF LARGE-SCALE COMPRESSIBLE CONVECTION</li> <li>o CALIBRATE STELLAR CONVECTION THEORIES</li> </ul>
<u>AMPLITUDE AND WIDTH OF PEAKS IN POWER SPECTRA</u>	DETERMINE MODE LIFETIMES AND QUALITY OF RESONANT CAVITIES <ul style="list-style-type: none"> <li>o EXAMINE MODE EXCITATION AND DAMPING</li> </ul>

### 3. NATURE OF SOLAR OSCILLATIONS

#### a. Initial Discovery of 5-Minute Oscillations

Solar seismology encompasses a variety of acoustic and gravity waves possessing a broad range of periods. The 5-minute acoustic oscillations have received particular attention because they are the ones most easily observed. The subject of solar oscillations really began with the first detection of both 5-minute oscillations and supergranulation, as reported by Leighton (1960) and Leighton, Noyes and Simon (1962). This led to many papers on the quasi-periodic variations seen in Doppler velocities over a range of heights in the photosphere. In particular, Doppler observations carried out in one spatial dimension (such as along the spectroscope slit) and in time revealed that the 5-minute oscillations looked like wave packets which lasted through about 6 or 7 periods and had spatial scales of coherence of about 30 Mm. An example of such observations is shown in Figure 1. The 5-minute oscillations there have peak Doppler velocities of about  $1000 \text{ ms}^{-1}$ , thus making them one of the more energetic velocity signals in the photosphere, on a par with those of granulation and supergranulation. Indeed, the vigorous 5-minute oscillations largely mask the persistent flows of supergranulation and mesogranulation unless one filters out their effects by some form of averaging in time.



**Figure 1.** Doppler observation of 5-minute oscillations on the Sun, as measured in one spatial dimension and time, reveals what appear to be wave packets in both space and time. When observed at a given site in the solar atmosphere, such quasi-periodic oscillations attain peak velocity amplitudes of about  $1000 \text{ ms}^{-1}$ . These signals are actually the superposition of nearly  $10^7$  acoustic resonance modes of oscillation, each with amplitudes of order  $20$  to  $50 \text{ cm s}^{-1}$ , that interfere both constructively and destructively to yield these apparent wave packets.



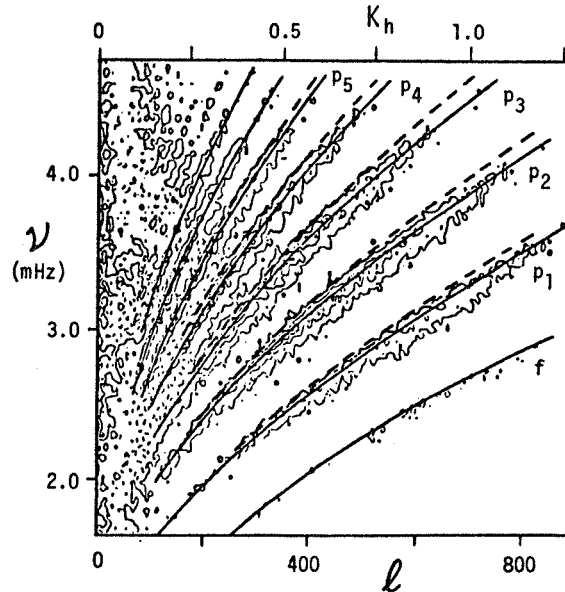
Given the wave packet character seen in Figure 1, much of the early interpretative work concentrated on the possibility that the 5-minute oscillations were resonances set up in the solar atmosphere by the emergence of new granulation convection cells. Since granules have typical lifetimes of about 10 minutes, it seemed reasonable that acoustic waves with a range of similar periods could be excited in the stable atmosphere by it being thumped from below by overshooting convective elements. Although variations of this proposition fostered many papers, the explanation lay elsewhere.

#### **b. Resonant Acoustic Cavity**

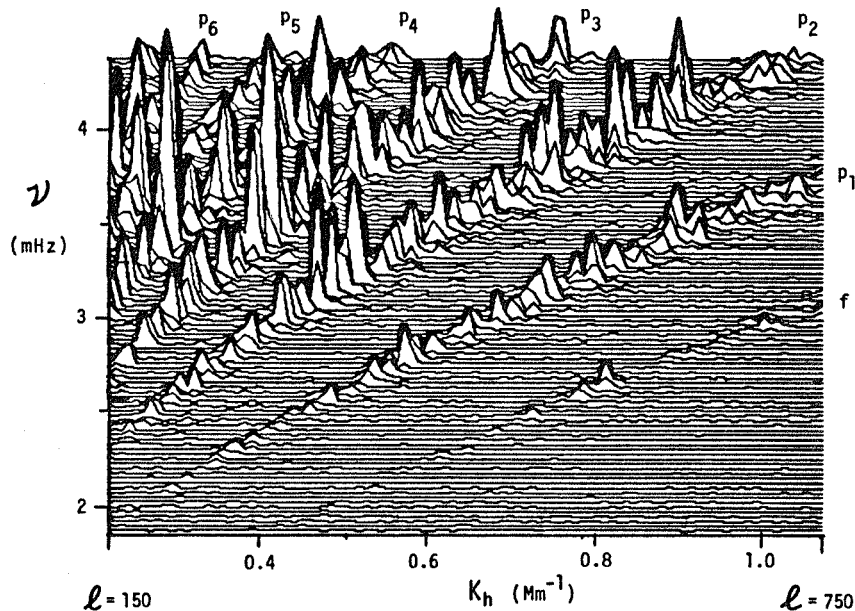
The existence of an acoustic resonant wave cavity below the solar surface was first proposed by Ulrich (1970) and by Leibacher and Stein (1971) to explain the 5-minute oscillations. Ulrich further made the important prediction that the depth of such an effective cavity would vary with the horizontal wavenumber  $k_h$  (or associated spherical harmonic degree  $\ell$ ) of the acoustic wave, with increasing  $k_h$  leading to shallower cavities. The cavity would respond most strongly at a number of discrete values of frequency  $\nu$  for each value of  $k_h$ . Power in the oscillations would thus be expected to lie preferentially along multiple curves in  $k_h$ - $\nu$  space. That such characteristic ridges in power actually exist when observing the real Sun was first demonstrated by Deubner (1975) and refined by other observations that followed (e.g. Rhodes, Ulrich and Simon 1977; Deubner, Ulrich and Rhodes 1979; Hill, Toomre and November 1983).

The observational breakthrough came from observing solar Doppler velocities over a two-dimensional spatial window, sampling an array of say  $128 \times 512$  spatial points with 2" resolution, with the longer dimension covering about half the diameter of the Sun. Further, such Doppler observations should be repeated about once per minute for an overall time interval of about 6 hours or longer. By spatially averaging the data along the narrower dimension, one can then isolate the resonant acoustic waves traveling around the Sun on one great circle from most others. Fourier transforms applied to the data in the remaining spatial dimension and in time then yield power spectra in  $k_h$  and  $\nu$ , much as depicted in Figure 2. Power is there concentrated along specific loci in the  $k_h$ - $\nu$  plane, in reasonable conformity with frequencies obtained from recent theoretical calculations. That these are ridges in power is further emphasized in the perspective view of another  $k_h$ - $\nu$  diagram in Figure 3 obtained from Doppler observations, which also serves to indicate that the distribution of power along the ridges involves fairly jagged structures.

The Doppler signals as observed in Figure 1 are not the result of discrete excitations that set only nearby portions of the atmosphere into oscillation, with an amplitude of say  $1000 \text{ ms}^{-1}$ . Rather, over  $10^7$  acoustic wave modes are involved, each with amplitudes of order 20 to 50  $\text{cm s}^{-1}$ . These normal modes of oscillation may be excited by stochastic interaction with the turbulence in the convection zone. Excitation and damping of a given wave mode would proceed almost continuously in a spatially distributed fashion. It may also be that the modes are



**Figure 2.** Power spectrum in frequency  $\nu$  and horizontal wavenumber  $k_h$  (or degree  $l$ ) from Doppler observations of high-degree 5-minute oscillations. Superposed on these ridges in power are theoretical predictions for frequencies of oscillation modes obtained from a solar model with a helium content of 25% by mass (shown as solid curves) and another which is helium deficient (19%, shown dashed). These curves are labelled by  $f$  to identify the loci of frequencies of the fundamental (or surface gravity) modes and by the values of the radial order  $n$  of the acoustic modes.



**Figure 3.** Perspective view in  $\nu$  and  $k_h$  of power spectrum of high-degree acoustic oscillations observed in Doppler shifts of a Mg I line. The considerable variability in power along the ridges is in part caused by noise introduced by effects of atmospheric seeing and in part by mode beating.

self-excited by something like the  $\kappa$ -mechanism which permits them to extract energy from the radiation field. These  $10^7$  modes interfere both constructively and destructively to produce the overall patterns of oscillatory velocities that are observed. The apparent lifetimes of the wave packets in Figure 1 correspond roughly to the beating period between modes, for instance from adjacent ridges, and the apparent spatial coherence lengths similarly correspond to the spatial beating scale between such modes.

### c. Notation for Normal Modes of Resonance

Oscillations within a spherical object like the Sun can be represented in terms of a superposition of many normal modes, each of which varies sinusoidally in time  $t$  with frequency  $\nu$ . The spatial structure of the wave pattern associated with any normal mode is most naturally described in spherical polar coordinates  $(r, \theta, \phi)$  as the product of a function of radius  $r$  and a function of colatitude  $\theta$  and longitude  $\phi$ . This permits one to express the radial component of velocity in a specific normal mode as the real part of

$$v(r, \theta, \phi, t) = V_n(r) Y_\ell^m(\theta, \phi) e^{-2\pi i \nu t}, \quad (1)$$

where  $Y_\ell^m \equiv P_\ell^m(\cos\theta) e^{im\phi}$  is a spherical harmonic function of degree  $\ell$  and order  $m$  involving the associated Legendre function  $P_\ell^m$ . For each pair  $\ell$  and  $m$ , there is a discrete spectrum of modes with distinct frequencies  $\nu$ . These possess differing spatial structures with radius, as represented by the eigenfunctions  $V_n(r)$  which are oscillatory in space and typically possess  $n$  zeros or nodes. The angular degree  $\ell$  measures the total horizontal component of the wavenumber,  $k_h$ , with

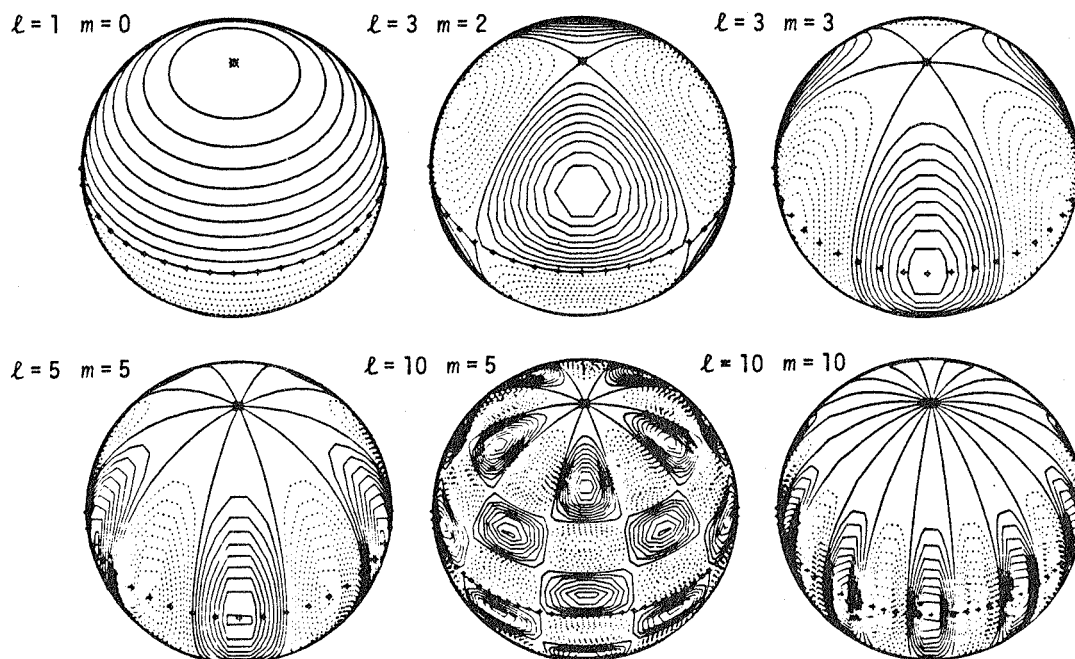
$$k_h = [\ell(\ell+1)]^{1/2}/r \quad (2)$$

at radius  $r$ . The azimuthal order  $m$  measures the component of  $k_h$  around the axis of the spherical coordinate system, and is constrained to take the  $2\ell+1$  values  $-\ell, \dots, +\ell$ . Although the effective radial component  $k_r$  of the wavenumber varies with depth (with  $k_r$  vanishing at the turning points of the cavity and becoming imaginary outside), it is convenient to regard  $n\pi/R$  as a mean radial wavenumber, where  $R$  is the radius of the Sun.

The frequencies  $\nu$  of the normal modes thus depend on what may be called the quantum numbers  $n$ ,  $\ell$ , and  $m$ , and we denote this by writing  $\nu \equiv \nu_{n, \ell, m}$ . One can measure the  $\ell$  and  $m$  values of the modes by making spatially resolved observations of the disturbances at the solar surface. However, the  $n$  value of a mode cannot be measured directly, since it depends on unobservable variation with depth. The complete identification of a mode will therefore require some theoretical knowledge of how  $\nu$  varies with the three quantum numbers. Before we turn to that in the next section, we can deduce that if the Sun were to be perfectly spherically symmetric,  $\nu$  must be independent of  $m$ . This comes about because the azimuthal order  $m$  depends on the choice of axis of the coordinate system, yet in a spherically symmetric configuration

all choices must be equivalent. Such a degeneracy of  $\nu$  on  $m$  can be relaxed by any departure from spherical symmetry, thereby resulting in fine structure in the frequency spectrum that is  $m$  dependent. The breaking of symmetry can be achieved by rotation or magnetic fields, and thus measurements of the fine structure permit deductions about such phenomena within the Sun.

The appearance of some spherical harmonics  $Y_{\ell}^m$ , as projected onto the solar disk, is shown in Figure 4. At any instant in time, the pattern of Doppler velocities of an individual oscillation mode would thus consist of regions of approaching flow (shown by solid contours) alternating with receding ones (dotted contours). Increasing  $\ell$  or  $m$  yields increasingly complex spatial patterns for the velocities. For zonal modes (with  $m = 0$ ), the nodal lines of zero velocity are lines of latitude; for sectoral modes (with  $m = \ell$ ), they are lines of longitude. Within this subject of solar seismology, the oscillations have further come to be grouped into three classes according to their angular degree: **low-degree** modes for  $\ell \leq 3$ , **intermediate-degree** modes for  $4 \leq \ell \leq 100$ , and **high-degree** modes for  $100 \leq \ell \leq 1000$ .



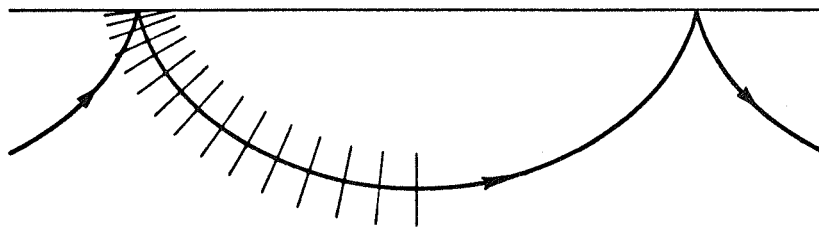
**Figure 4.** Contours of Doppler velocities for selected normal modes of oscillation as projected onto the solar disk in the absence of any other modes. At any instant in time, the patterns consist of approaching flow (shown as solid contours) alternating with receding ones (dotted). Increasing the degree  $\ell$  or azimuthal order  $m$  of such spherical harmonics  $Y_{\ell}^m$  produces more complex spatial patterns. Shown also are a zonal mode ( $m = 0$ ) which is axisymmetric, and a sectoral mode ( $m = \ell$ ) which is confined about an equatorial strip.

An interesting feature of the sectoral modes is that with increasing degree they become more confined about an equatorial strip bounded by the latitudes  $\cos^{-1}\{m/[\ell(\ell+1)]^{1/2}\}$  (e.g. Gough 1983). This comes about because sectoral modes at high  $\ell$  are much like plane-parallel waves, but for geometrical problems in fitting onto a spherical surface; their lateral confinement results from destructive interference at higher latitudes where their apparent phase velocities with longitude are proportionately greater than at the equator. Such a property of sectoral modes was exploited in obtaining the high  $\ell$  power spectra shown in Figures 2 and 3: a long but narrow observing window aligned with the solar equator serves as a reasonably sensitive detector of sectoral modes to the exclusion of others. Of course, any great circle on a sphere, not just the equator, should possess waves that are propagating along it, and they too can be described most simply in terms of sectoral modes if suitable rotations are made for the spherical coordinate system so that its equator coincides with that great circle.

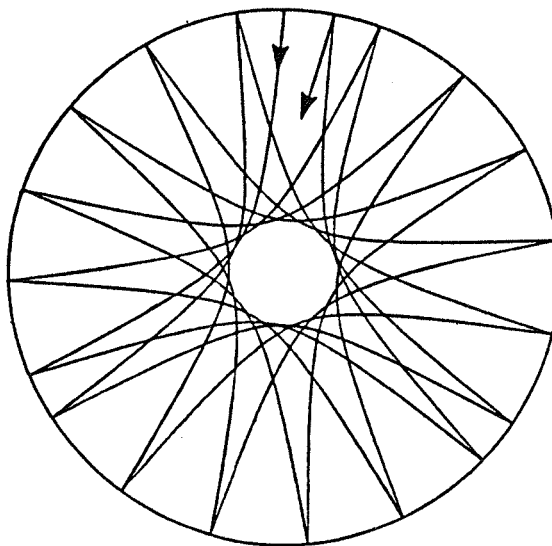
Figure 4 also serves to explain why the sensitivity of integrated sunlight observations of the oscillations are limited to low-degree modes with  $\ell \lesssim 3$ . In such observations made without imaging, one records only an average of the velocity signal over much or all of the solar disk, and since contributions from regions with velocities of opposite sign cancel, modes of high degree yield little as a net signal. In the absence of spatial resolution, the modes that survive in the integrated signal can be identified solely by comparing their frequencies with theory and their relative amplitudes with models of the response of a given observing aperture to specific spherical harmonics.

#### d. Properties of the Acoustic Modes

We shall now turn to reconsider how the resonant wave guide works for the originally discovered 5-minute acoustic modes. This turns out to be particularly simple because the wavelengths of those modes are short compared to the radius of the Sun, and thus spherical curvature effects may be ignored. These modes are all of high angular degree  $\ell$  ( $\ell \gtrsim 100$ ) and low radial order ( $n \lesssim 10$ ). The original observations concentrated on high-degree sectoral modes, and since these can be viewed as nearly plane-parallel waves, they are usually distinguished by  $k_h$  in place of  $\ell$ . Turning to Figure 5 which sketches ray paths and surfaces of constant phase for such acoustic waves, one finds that as the waves approach the surface of the Sun from below they are reflected back downward by the very steep density gradient there. As a given wave then begins to propagate downward, it is progressively refracted around until it is once more headed upward. This comes about because the deep portions of the wavefront are traveling at a faster sound speed  $c$  (recalling that  $c \propto T^{1/2}$ , where  $T$  is the mean temperature which increases with depth). At the point of lower reflection, the wave is propagating horizontally, and reflection occurs at the depth  $d$  into the Sun where  $c = 2\pi\nu/k_h$ . Therefore for a given frequency  $\nu$ , waves of greater horizontal wavenumber  $k_h$  are refracted faster and penetrate less deeply since their wavefronts (as in Figure 5) are more tilted from the horizontal. The depth of penetration  $d$  of the wave into the Sun, or



**Figure 5.** Typical ray paths and surfaces of constant phase for an acoustic mode of high degree  $\ell$  confined to a shallow region just below the solar surface (shown as the horizontal line). An acoustic wave is trapped between an upper turning point just beneath the photosphere where the wave is reflected by the rapid decrease in the density, and a lower turning point where the wave is turned around by the increase in sound speed with depth.



**Figure 6.** Global nature of low-degree acoustic modes is shown in the multiply-reflected ray paths associated with this mode of high radial order ( $n/\ell \approx 5$ ) which propagates throughout most of the solar interior. The circle represents the solar surface.

thus the effective depth of the particular acoustic cavity or wave guide, varies as  $d \approx (2n+3)/k_h$ . Although all these 5-minute acoustic modes are evanescent (non-propagating) in the atmosphere, they are able to tunnel into that region with sufficient amplitude to be readily observable.

Working out the travel time between reflections in such resonant trapping of acoustic modes within the convection zone leads to a dispersion relation approximately of the form

$$v \approx [(2n+3) g k_h]^{1/2}, \quad (3)$$

where  $g$  is the acceleration due to gravity. The precise relation depends sensitively upon the stratification below the surface. Such a relation defines the curves shown in Figure 2, or the centroids of the ridges in Figure 3, with the order  $n$  distinguishing one curve or ridge from another. Thus such specific groups of acoustic modes are identified by their order  $n$ , say as  $p_2$  when  $n = 2$ . Such ridge classification ignores the possible existence of chromospheric modes which contribute additional nodes to the radial eigenfunctions. The lowest ridge in those figures corresponds to a surface gravity mode, usually called the  $f$  mode, which has  $n = 0$ .

The plane-parallel description of the oscillations, as in Figure 5, is inadequate when  $k_h$  is small, since then the horizontal wavelength  $2\pi/k_h$  and the penetration depth  $d$  are significant fractions of the solar radius. Therefore effects of curvature must explicitly be taken into account. Further, the horizontal phase speed of the waves increases with decreasing  $k_h$ , in essence because the waves penetrate deeper and experience a proportionately higher sound speed. Indeed, for  $k_h$  less than about  $0.15 \text{ Mm}^{-1}$ , modes propagate around the Sun in less than about 10 hours, which is also less than their estimated lifetimes. Thus one would expect that only those waves survive which interfere constructively around the solar circumference, thereby selecting particular values of  $k_h$ . This suggests that oscillation modes of low  $k_h$  are really global in character, and they are most readily described in terms of spherical harmonics  $Y_l^m$ .

The global nature of low-degree acoustic modes is also apparent in their ray paths within the Sun. Figure 6 shows the path of a typical multiply-reflected ray contributing to an acoustic mode of high radial order  $n$ . Such a mode, with  $l \ll n$ , propagates throughout most of the solar interior; only a central region is avoided. A normal mode may be regarded as the interference pattern of disturbances between many traveling waves, each possessing a ray path of the form shown. Incidentally, those ray paths do not generally close back on themselves after multiple bounces, for it is only the overall interference pattern that possesses spatial periodicities in the form of spherical harmonics.

The dispersion relation for acoustic modes can be obtained by insisting that information propagating along the ray paths is in phase with the interference pattern. The relation (3) for high-degree modes

(with  $\ell \gg n$ ) was so deduced. Assuming that the Sun is spherically symmetric, similar reasoning applied to low-degree modes (with  $\ell \ll n$ ) leads to the dispersion relation (e.g. Gough 1983)

$$\nu_{n,\ell} = \nu_0 \left( n + \frac{1}{2}\ell + \delta \right) + \epsilon_{n,\ell} , \quad (4)$$

where

$$\nu_0 = \left[ 2 \int_0^R \frac{dr}{c} \right]^{-1} \quad (5)$$

is the reciprocal of the sound travel time from the surface at radius  $R$  to the core and back, and  $\delta$  is a small term which is related to the effective polytropic index of the outer layers of the solar envelope. The next term in the expansion,

$$\epsilon_{n,\ell} \approx - \nu_0 \frac{\alpha \ell(\ell+1) - \beta}{n + \frac{1}{2}\ell + \delta} , \quad (6)$$

involves the positive dimensionless constants  $\alpha$  and  $\beta$  which are of order unity. Such asymptotic relations (see also Vandakurov 1967, Tassoul 1980) imply that modes of given degree  $\ell$  but successive order  $n$  occur at discrete frequencies approximately evenly spaced by  $\nu_0$ , with  $\nu_0$  expected to be about 136  $\mu\text{Hz}$  for the Sun. Further, increasing  $\ell$  by 2 and decreasing  $n$  by 1 leaves  $\nu$  almost unchanged, since for  $n \gg \ell$ ,  $\epsilon_{n,\ell}$  provides only a small correction to the leading term. Thus arise groups of uniformly spaced frequencies belonging alternately to modes of odd and even degree, with a spacing  $\nu_0/2$  between groups.

#### e. Properties of the Gravity Modes

Resonant cavities also exist for internal gravity waves in the Sun. Such waves can propagate only where the stratification is stable to convection, and further where their frequency  $\nu$  is less than the local buoyancy (or Brunt-Vaisala) frequency  $N$ . Thus propagation is possible in the solar interior below the convection zone, and also within the atmosphere. The upper and lower reflection points of any given cavity correspond to where  $N$  has approached  $\nu$ , aside from some factors depending on the degree  $\ell$  of the wave. The  $g$  modes of the deep interior all share an upper reflection point near the base of the convection zone; their amplitudes decay throughout that zone where the modes are evanescent. Since the decay rate increases with  $\ell$ , only modes of low degree are likely to be detectable in the atmosphere.

Theoretical calculations of dispersion relations for gravity waves reveal that these normal modes should be nearly uniformly spaced in period  $P_{n,\ell}$ , unlike  $p$  modes which are nearly uniformly spaced in



frequency. These g modes should possess periods

$$P_{n,\ell} \approx \frac{n + \frac{1}{2}\ell + \gamma}{[\ell(\ell+1)]^{1/2}} P_0 \quad (7)$$

with

$$P_0 = 2\pi^2 \left[ \int_0^{r_c} N \frac{dr}{r} \right]^{-1}, \quad (8)$$

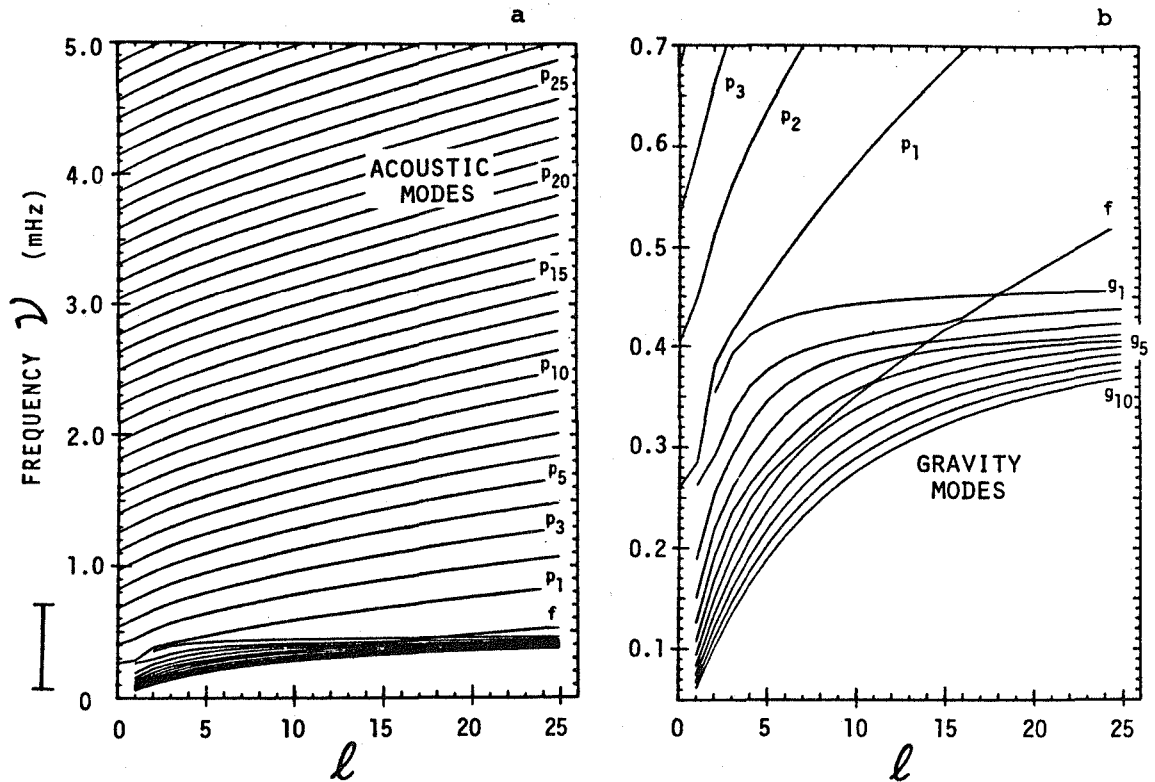
where  $\gamma$  is a term that depends on the structure just below the base of the convection zone, with the latter at radius  $r_c$ . The buoyancy frequency  $N$  is given by

$$N^2 = g \left[ \frac{1}{\Gamma_1} \frac{d \ln p}{dr} - \frac{d \ln \rho}{dr} \right], \quad (9)$$

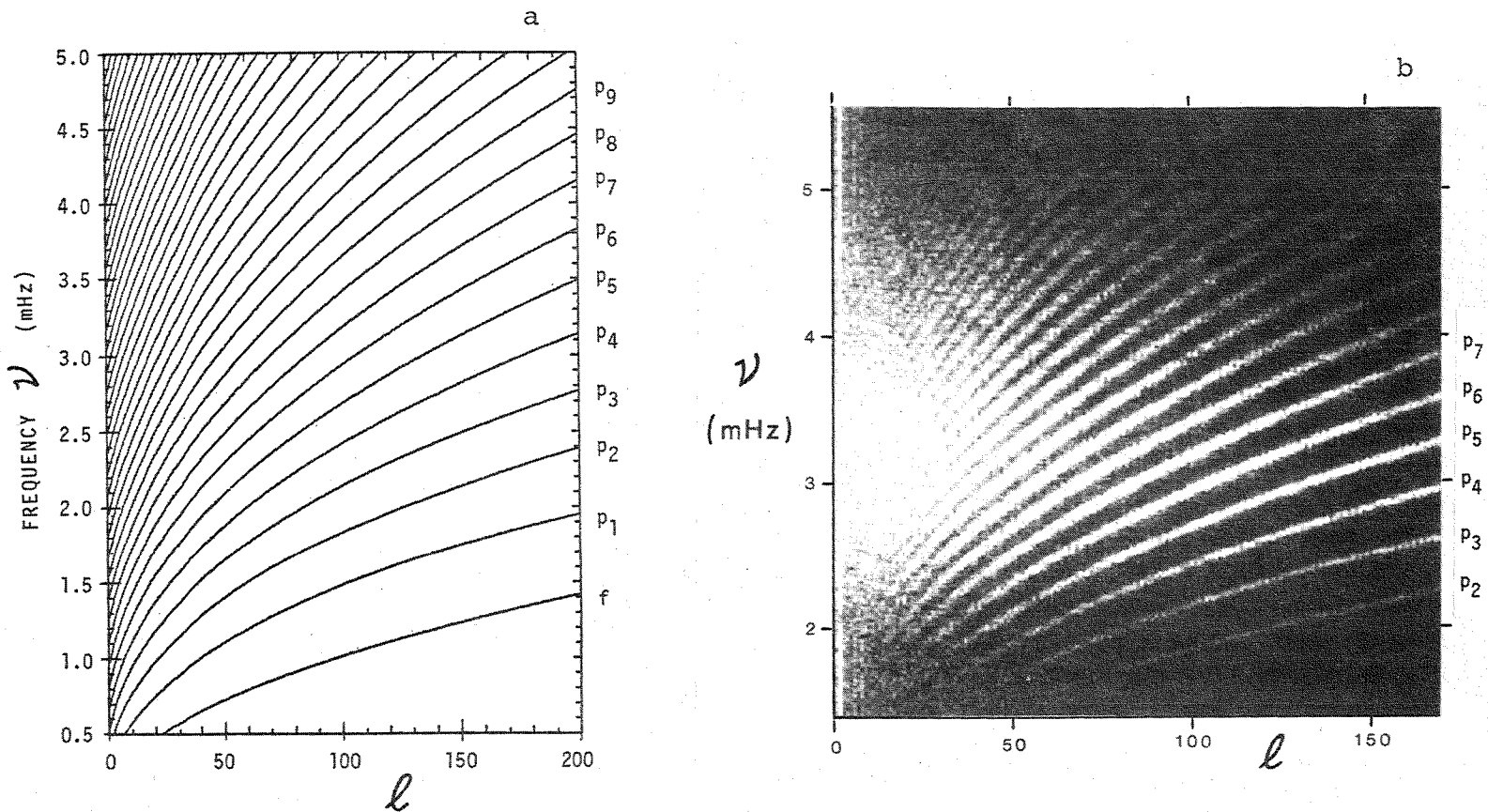
where  $p$  is pressure,  $\rho$  is density and  $\Gamma_1 = (\partial \ln p / \partial \ln \rho)_s$ ,  $s$  being specific entropy. Having access to precise frequencies of g modes would thus enable one to sample rather different properties of the stratification than would be afforded by p modes and their sampling of an integral of the sound speed  $c$ .

#### f. Frequencies and Radial Structure of Oscillations Obtained from Numerical Models

Figure 7 shows the frequencies  $\nu_{n,\ell}$  obtained from a solar model for sequences of acoustic and gravity modes of low and intermediate degree. These theoretical calculations, like all others to date in the subject, have used linearized equations for the dynamics of the oscillations, which is warranted given the small observed amplitudes for individual modes. Such calculations (e.g. Ando and Osaki 1977; Christensen-Dalsgaard and Gough 1981; Scuflaire, Gabriel and Noels 1982; Ulrich and Rhodes 1983) incorporate nonadiabatic effects in the outer layers of the Sun, but generally ignore effects of mode excitation and dissipation by the turbulence in the convection zone. Thus although the frequencies of the oscillations are likely to be determined quite accurately for a given solar model, issues of growth rates for the modes are quite uncertain (cf. Goldreich and Keeley 1977). Figure 7 displays loci of frequencies at fixed  $n$  as  $\ell$  varies for both p and g modes; these curves are designated by their radial order  $n$ . The frequencies  $\nu_{n,\ell}$  are actually determined only at integral values of  $\ell$ , given the global nature of the modes, but these discrete frequencies have been joined by smooth curves for clarity. Figure 7a illustrates the nearly uniform spacing in frequencies for p modes with  $n$  at fixed  $\ell$ , consistent with the asymptotic relation (4). It also shows the crowding of g modes at much lower frequencies or longer periods. The portion of the low frequency domain delineated by the vertical bar is expanded in Figure 7b to reveal the frequency dependence of a selection of g modes, terminating arbitrarily with  $g_{10}$ . As predicted by the asymptotic relation (7) for such g modes, the frequencies get increasingly crowded both with increasing  $n$  and  $\ell$ . Such close spacing of the g modes in frequency places severe demands on the frequency resolution that must be



**Figure 7.** Frequencies  $\nu_{n,l}$  obtained from a theoretical solar model for p and g modes of low and intermediate degree  $l$ . Displayed are loci of frequencies at fixed radial order  $n$  as  $l$  is varied, with the curves identified by  $n$ . Shown in (a) is the frequency range of 0 to 5 mHz, and in (b) an expanded view of the lower frequencies from 0.05 mHz (period of about 333 min) to 0.70 mHz (period 24 min). The left panel (a) shows that the p modes possess nearly uniform spacing in frequencies at fixed  $l$  as  $n$  is varied, provided  $n \gg l$ ; the crowding of the g modes at low frequencies makes them nearly indistinguishable. The right panel (b) shows the gravity modes  $g_1$  to  $g_{10}$  far more clearly, and evidently the frequencies converge with increasing  $l$ . For these g modes at fixed  $l$ , it is the spacing in period and not in frequency which is nearly uniform as  $n$  is varied, though there are evident perturbations here due to avoided crossings of frequencies.



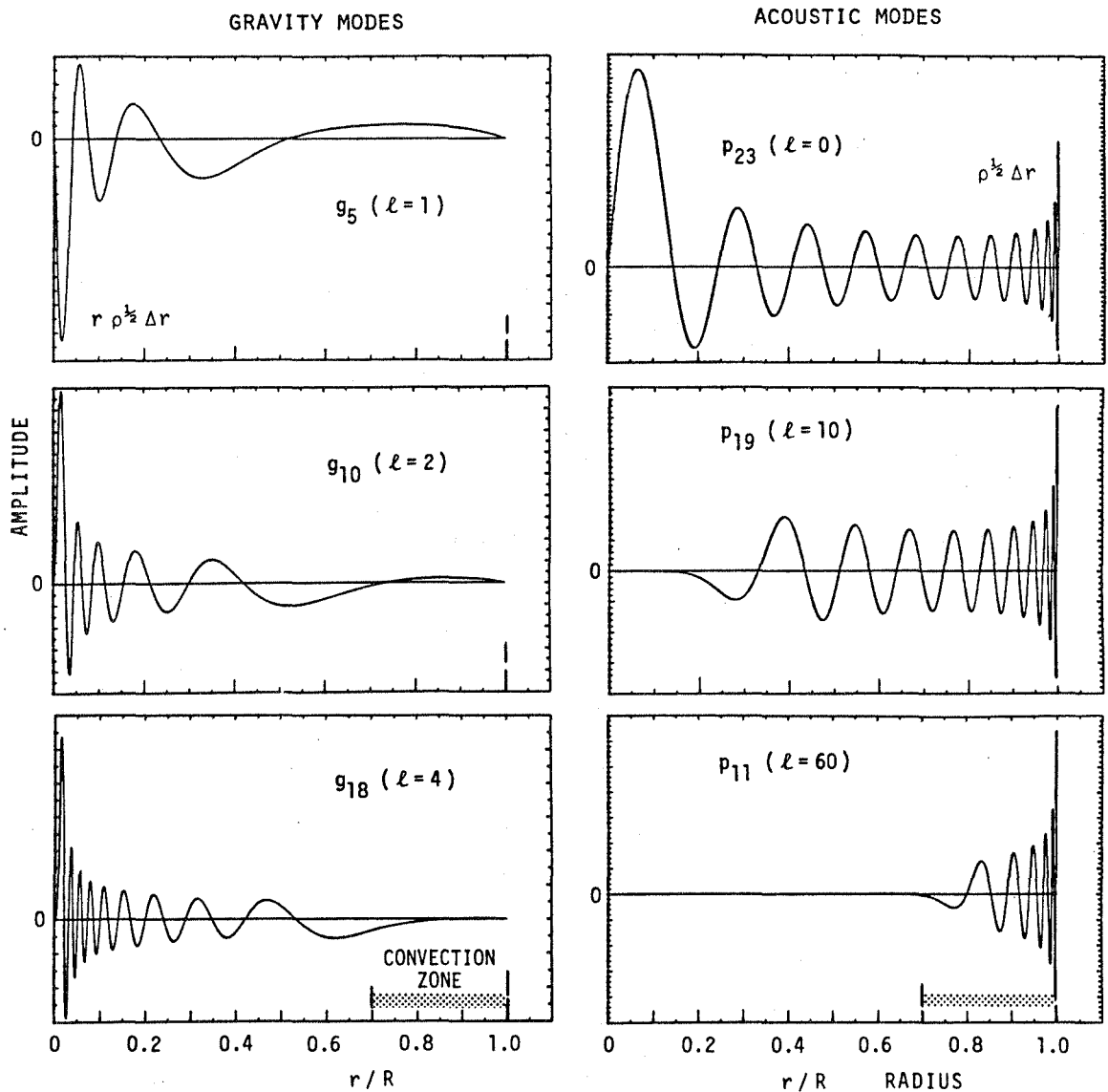
**Figure 8.** Comparison of frequencies of acoustic modes, primarily of intermediate degree  $l$ , as obtained from theory and observation. Shown in (a) are loci of frequencies determined from a theoretical solar model, and in (b) a power spectrum in  $\nu$ - $l$  obtained from Doppler observations with imaging of the zonal modes of oscillation of the Sun. The striking fan structure of frequencies in these observations has provided an unambiguous link between acoustic modes of high and low degree, thereby clearly identifying the radial order  $n$  in low  $l$  observations; a knowledge of  $n$  is vital for detailed modelling of the interior stratification.

achieved in observational power spectra if the modal identification is to be unambiguous.

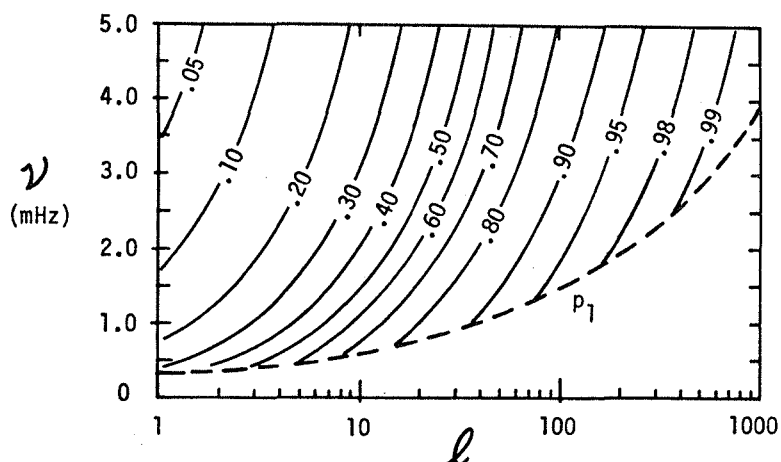
Given the limited range in  $\ell$ , Figure 7 provides only a partial view of the frequencies of the intermediate-degree modes. This is rectified by Figure 8a which displays theoretical p mode frequencies for  $\ell$  extending up to 200, thereby also showing frequencies for some of the high-degree modes. This plot emphasizes that at higher  $\ell$  the frequency interval between modes of differing  $n$  is considerably wider than at low  $\ell$ . The reality of such a large fan structure of frequencies is confirmed by the remarkable observational power spectrum, shown in Figure 8b, which was obtained by Duvall and Harvey (1983) when observing intermediate-degree modes. This power spectrum also makes a clear link between the ridges readily distinguished by radial order  $n$  at high  $\ell$  and the modes at low  $\ell$  where the  $n$  identification was previously unclear.

The plots of  $\nu_{n,\ell}$  also remind us that nearly identical temporal frequencies can be achieved with a variety of normal modes, each mode sampling quite different regions within the Sun. Figure 9 illustrates this by showing the variation with proportional radius  $r/R$  of scaled displacement amplitudes for several p and g modes. The three examples of acoustic modes on the right have similar frequencies, of 3.3 mHz (or periods of about 5 min), while the gravity modes on the left all have frequencies of 0.10 mHz (periods of about 165 min). For the p modes, increasing the degree  $\ell$  leads to the modes becoming increasingly confined to a region close to the surface; note that similar frequencies are achieved in this example by suitably decreasing the radial order  $n$  as  $\ell$  is increased. For the g modes, increasing  $\ell$  and  $n$  yields more concentrated and intricate modal structure near the core of the Sun. Although these global g modes are confined to the stable region beneath the convection zone as far as their energy is concerned, the amplitudes of the low  $\ell$  modes decay slowly enough in tunneling through the convection zone that they may have measurable displacements or velocities in the solar atmosphere.

As Figure 9 readily shows, these g modes have the greatest potential of sampling the stratification, as  $N$ , near the core, though the low-degree p modes also penetrate quite deeply into the Sun and can sample the variation of sound speed  $c$  there. The extent of penetration of an acoustic mode can be measured by the lower reflection or turning point of the mode, and this depends both upon  $\nu$  and  $\ell$ . Figure 10 shows how that turning point, plotted as proportional radius  $r/R$ , varies with  $\nu$  and  $\ell$ . It is evident that p modes of low degree probe more deeply than those of high degree. Thus although p-modes of lower degree and higher frequency sample conditions over greater ranges of depth, scientific inferences of conditions at great depths will be most reliable if modes of high degree are measured accurately as well. This comes about because all p modes are very sensitive to conditions near the surface, as can be judged from their amplitudes in Figure 9. Effects of conditions at considerable depths can be separated out best if accompanied by precise measurements of high-degree modes confined near the surface. This is an issue of vital importance in assessing



**Figure 9.** Scaled displacement amplitudes with proportional radius  $r/R$  of some acoustic and gravity modes within the Sun as obtained from a theoretical model. The theoretical eigenfunctions for the radial displacements, say  $\Delta r$ , have been scaled by the indicated factors of density  $\rho$  and radius  $r$  for clarity of display. The three acoustic modes shown on the right all possess similar frequencies of 3.3 mHz (periods of about 5 min), obtained by decreasing the radial order  $n$  while increasing degree  $\ell$ . These p modes are increasingly confined to a region close to the surface (at  $r/R = 1$ ) as  $\ell$  is increased. The three gravity modes on the left all have frequencies of 0.10 mHz (periods of 165 min). Increasing  $\ell$  and  $n$  yields g modes with more complicated structure near the core of the Sun. All these g modes possess an upper reflection point near the base of the convection zone, with their amplitudes decaying throughout that zone where the modes are evanescent.



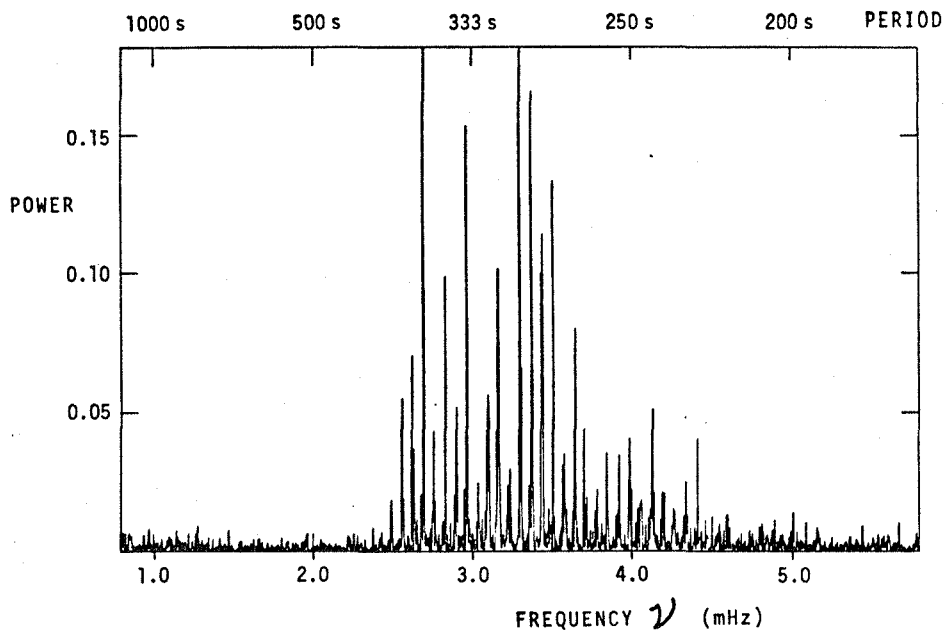
**Figure 10.** Contours showing the extent of penetration of acoustic modes into the Sun as a function of degree  $\ell$  and frequency  $\nu$ . The contours are labelled by the proportional radius of the lower turning point of a mode. For reference, the curve identified as  $p_1$  marks the lowest frequencies attainable by acoustic modes as a function of  $\ell$ .

benefits of observations to be made from space, for the necessary high  $\ell$  measurements are particularly impacted by seeing distortions within the Earth's atmosphere.

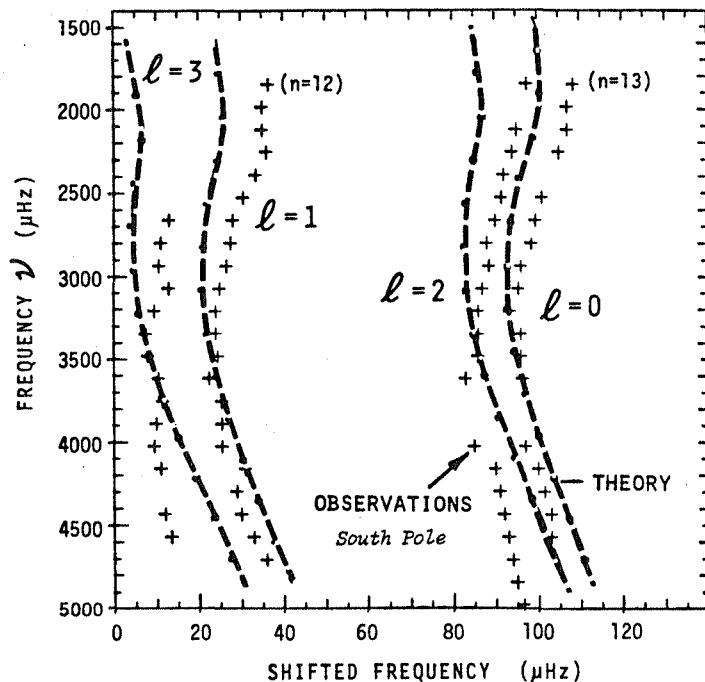
#### g. Comparison of Observed and Computed Frequency Spectra of Low-Degree Acoustic Modes

Observations of low-degree p modes using sodium or potassium resonance cells as Doppler detectors have yielded unexcelled frequency spectra to date (e.g. Grec, Fossat and Pomerantz 1980; Claverie et al. 1980). The potential of seismology was convincingly demonstrated by observations carried out from the South Pole during austral summer to avoid day-night gaps in the data string. Figure 11 shows the power spectrum obtained from such observations during a continuous 5-day time interval. Given the spatial response to various  $Y_\ell^m$  of such observations without imaging, the succession of peaks in frequency correspond to modes  $\ell = 0, 1, 2$  and  $3$  of varying degree  $n$  (cf. Christensen-Dalsgaard and Gough 1982). The observed power in these low-degree acoustic modes is greatest in the frequency range 2.5 mHz to 4.5 mHz, which must be a consequence of preferential excitation in the convection zone, for Figure 7a reminds us that many p and g normal modes are feasible in principle at considerably lower and somewhat higher frequencies.

Detailed comparison of these observed frequencies with those obtained from solar models is facilitated by plotting the frequencies of peak power in a so-called echelle diagram in Figure 12, with that form suggested by the dispersion relation (4). The spectrum is divided into segments of length  $\nu_0 = 136 \mu\text{Hz}$  starting at an arbitrary frequency,



**Figure 11.** Power spectrum in frequency  $\nu$  of low-degree solar oscillations as observed in Doppler velocities without imaging from the South Pole during a continuous 5-day time interval.



**Figure 12.** Echelle diagram used to compare in detail the frequencies of oscillation modes (at the indicated low degrees  $l$  and successive radial orders  $n$ ) as obtained from a theoretical solar model and from full-disk Doppler observations carried out from the South Pole.

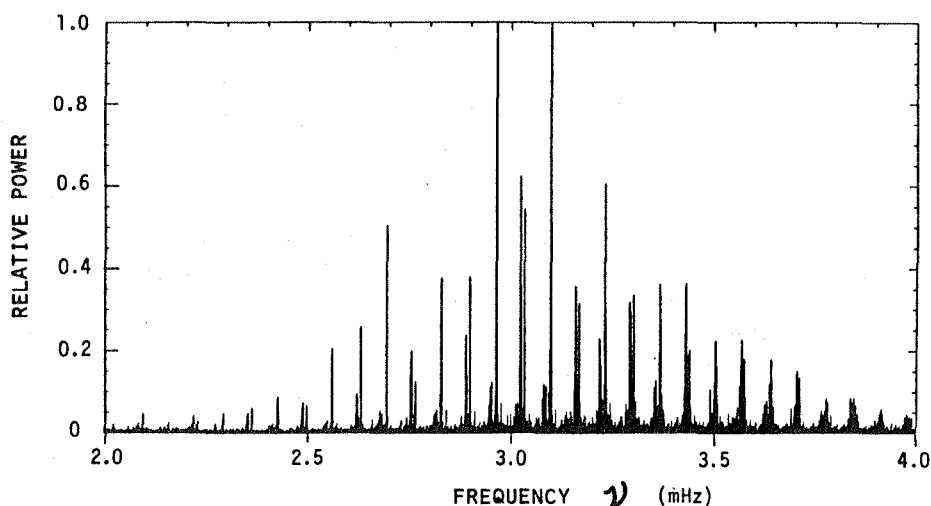
and the segments are presented directly beneath one another. That choice of  $\nu_0$  already places constraints on the stratification within solar models. Such a manner of display emphasizes that modes of odd and even degree fall into separate groups in frequency, and that successive segments at higher frequencies correspond to modes of successively higher radial order  $n$ . The separations in frequency between the  $\ell = 0$  and 2 and the  $\ell = 1$  and 3 modes are related to the  $\epsilon_{n,\ell}$  term of (6). The nearly vertical dashed curves show the loci of frequencies from a standard solar model, while the crosses denote frequencies from the South Pole observations.

The differences in frequency between observations and theoretical predictions are typically about 5  $\mu\text{Hz}$ , or about 0.2%, whereas the observational uncertainties are estimated to be about 2  $\mu\text{Hz}$ . By many standards this is remarkably good agreement, but the potential of seismological methods can only be realized if the predicted and observed frequencies match within the observational errors. Various standard solar models lead to slightly different sets of predicted frequencies, but none presently satisfies that stringent criterion. The differences in frequencies may involve issues of equations of state or opacities, effects on stratification by mixing below the convection zone or intermittent mixing of the core, elemental abundances and their distribution with radius, or possibly strong magnetic fields in the core. Mode excitation and decay may also contribute to the differences in ways that are not yet fully understood. Some of these processes are expected to influence frequencies at about the 1  $\mu\text{Hz}$  level. For that reason it is of essence to measure the frequencies to the level of 0.3  $\mu\text{Hz}$ , or to about 0.01%. Some low-degree  $p$  modes may maintain their coherence long enough for such observations to be feasible, as has been suggested by the promising observations of Isaak and colleagues obtained from a two-station network and shown as a power spectrum in Figure 13. Precise frequencies from these observations have not yet become available for detailed comparisons to be made with theoretical predictions. However, the potential quality of that spectrum promises to provide frequencies with a precision considerably better than 1  $\mu\text{Hz}$ . The strongest modes seen in full disk Doppler measurements have also been observed in integrated solar intensity by the Active Cavity Radiometer (ACRIM) on the Solar Maximum Mission satellite (e.g. Willson and Hudson 1981, Woodard and Hudson 1983). Such instruments provide a promising route for making complementary observations with high precision of intensity fluctuations of the low-degree modes.

#### **h. Detection of Gravity Modes**

The detection and identification of gravity modes present major challenges in seismology. It appears that the amplitudes in the solar atmosphere of all but the lowest-degree  $g$  modes may be below the sensitivity level of present detectors. Further, the inherently long periods of  $g$  modes require observations spanning years if suitable frequency resolution is to be achieved to distinguish the closely spaced frequencies of these modes. Observations of whole disk Doppler velocities from Crimea and Stanford have revealed an oscillation with a





**Figure 13.** Power spectrum in frequency  $\nu$  of low-degree solar oscillations obtained from full-disk Doppler observations carried out at two stations, one at Tenerife and another at Hawaii. The spectrum is based on observations spanning about 3 months in 1981, with as much as 22 hours of coverage per day.

period of 160.01 min, and if this is a solar  $g$  mode of high order, it appears to have maintained phase coherence over 7 years (e.g. Severny et al. 1976, Kotov et al. 1978, Scherrer et al. 1979, Scherrer and Wilcox 1983). The identification of that oscillation mode is still uncertain, but it may correspond to a  $g_{10}$  mode at  $\ell = 2$  (cf. Christensen-Dalsgaard and Gough 1976). Recent analysis of the 8 years of data from Crimea has yielded 32 frequencies with periods between 116 and 200 min which Severny et al. (1984) believe to be solar  $g$  modes and not terrestrial in origin.

There has been significant progress with techniques to mitigate effects of sidelobes and aliased peaks in power spectra due to daily gaps in these long data records (e.g. Scherrer, these proceedings). This has led to the tentative identification in the Stanford power spectra of a sequence of about a dozen peaks whose nearly uniform spacing in period satisfies relation (7), thereby suggesting that they correspond to  $g$  modes with periods in the range of 3 to 5 hours (Delache and Scherrer 1983, Scherrer 1984). They found the spacing in period, as in (8), to correspond to  $P_0 = 38.6 \pm 0.5$  min, which is rather greater than the values (33 to 36 min) to be expected in that frequency range from standard solar models calibrated by the  $p$  modes. Isaak et al. (1984) report having detected  $g$  modes in whole-disk Doppler measurements carried out by the Birmingham group (Figure 13 shows their power spectrum at the higher  $p$  mode frequencies). They find  $P_0 = 41.2$  min, inconsistent with the Stanford results. Further, Fröhlich and Delache (1984, and these proceedings) report statistical evidence from analyzing the ACRIM intensity data for  $P_0 = 45.0 \pm 0.2$  min, though those

oscillations are in the period range of 3.5 to 30 hours and thus not in direct conflict with the other values of  $P_0$ . The sources of discrepancy in  $P_0$  between the Stanford and Birmingham reports are unclear, but may arise from uncertainties in identifying the radial orders  $n$  or degrees  $\ell$  that correspond to the peaks (Gough 1984, Gabriel 1984).

The apparent differences in  $P_0$  between theory and observations possibly suggest that the buoyancy frequency  $N$  in the solar core should be reduced from what has been supposed in standard solar models. This might be accomplished by mild mixing of material between core and envelope to slightly reduce the gradient of chemical composition, and thus  $N$ , brought about by the nuclear reactions in the core (cf. Berthomieu et al. 1984). Certainly an accurate observational determination of  $P_0$ , which is an integral measure of internal stratification, will place important constraints on solar models, dealing both with the current state of the Sun and having implications on its evolutionary history. Such determinations would benefit greatly from observations of oscillations carried out with imaging Doppler analyzers to allow clear identification of the degree  $\ell$  of the  $g$  modes. Ground-based networks may be able to provide a significant component of such observations, though a space observatory with an operational lifetime of a number of years may prove to be essential for attaining the best signal-to-noise in determining the frequencies and relative amplitudes of the  $g$  modes.

#### 4. EFFECTS OF ROTATION IN SPLITTING THE FREQUENCIES

Any deviation from spherical symmetry can lift the degeneracy of frequency upon the azimuthal order  $m$  of the oscillation modes. In particular, rotation of the Sun will accomplish this, for then normal modes with different  $m$  will possess slightly different apparent frequencies  $\nu_{n,\ell,m}$ .

##### a. Simple Explanation of Splitting

One can understand how the frequency splitting comes about by turning to Figure 4, and concentrating on modes with the same  $\ell$  but differing  $m$ , as in the  $\ell = 10$ ,  $m = 5$  and  $10$  examples. What are shown there are instantaneous views of velocity patterns associated with specific propagating normal modes. A little later in time, each such pattern would appear as having been translated laterally (rotated around the polar axis) at the horizontal phase speed of that mode, simply because the underlying waves which produce such interference patterns are propagating around the Sun. Of course for given  $\ell$  and  $|m|$ , there is both a mode (say  $+m$ ) propagating eastward and another ( $-m$ ) propagating westward, both with identical angular phase speeds ( $2\pi\nu/m$ ) which vary inversely with  $m$ .

Consider now the simplest case where the rotation axis of the Sun coincides with the polar axis of the coordinate system, and further that its angular velocity  $\Omega(r)$  depends only on radius, with at first no latitudinal variation. The resulting advection of the wave patterns by

the rotation would cause the normal modes to appear to be swept past the viewer at different rates for the prograde (say +m) and retrograde (-m) modes. This could be interpreted as apparent changes in the temporal frequencies of the modes from the vantage point of the observer, for the frequency is a measure of the rate at which nodal lines proceed past a fixed reference point. That rate is a product of the apparent angular phase speed ( $2\pi\nu/m$ , augmented by a factor proportional to  $\Omega$ ) and the spatial wavenumber, here  $m$ , along the propagation direction. Thus the apparent shift in frequency is proportional to  $m\Omega$ , and thereby the degeneracy of frequency on  $m$  has been lifted. There are also modifications to oscillation frequencies from dynamical effects of Coriolis forces, but such changes are small for high-order  $p$  modes when compared to those of advection.

### b. Formal Estimates of Splitting

Given the simplifying assumption that the angular velocity  $\Omega(r)$  depends only on radius  $r$ , detailed calculations (e.g. Gough 1981, 1982b) of the effects of rotation show that a mode with frequency  $\nu_{n,\ell}$  is split into  $2\ell+1$  equally spaced components, which may be written as

$$\nu_{n,\ell,m} = \nu_{n,\ell,o} + \frac{m}{2\pi} \beta_{n,\ell} \int_0^R K_{n,\ell}(r) \Omega(r) dr . \quad (10)$$

The spacing of these components is determined by the integral over the volume of the Sun of the product of the angular velocity and a function that measures the sensitivity of the mode frequency to rotation at each depth. Such functions  $K_{n,\ell}$  are known as rotation splitting kernels, and examples of some are shown in Figure 14. The scaling factors  $\beta_{n,\ell}$  are of order unity, and are chosen to make the kernels  $K_{n,\ell}$  unimodular. For low-degree  $p$  modes with  $n \gg \ell$ , the frequency splitting is just

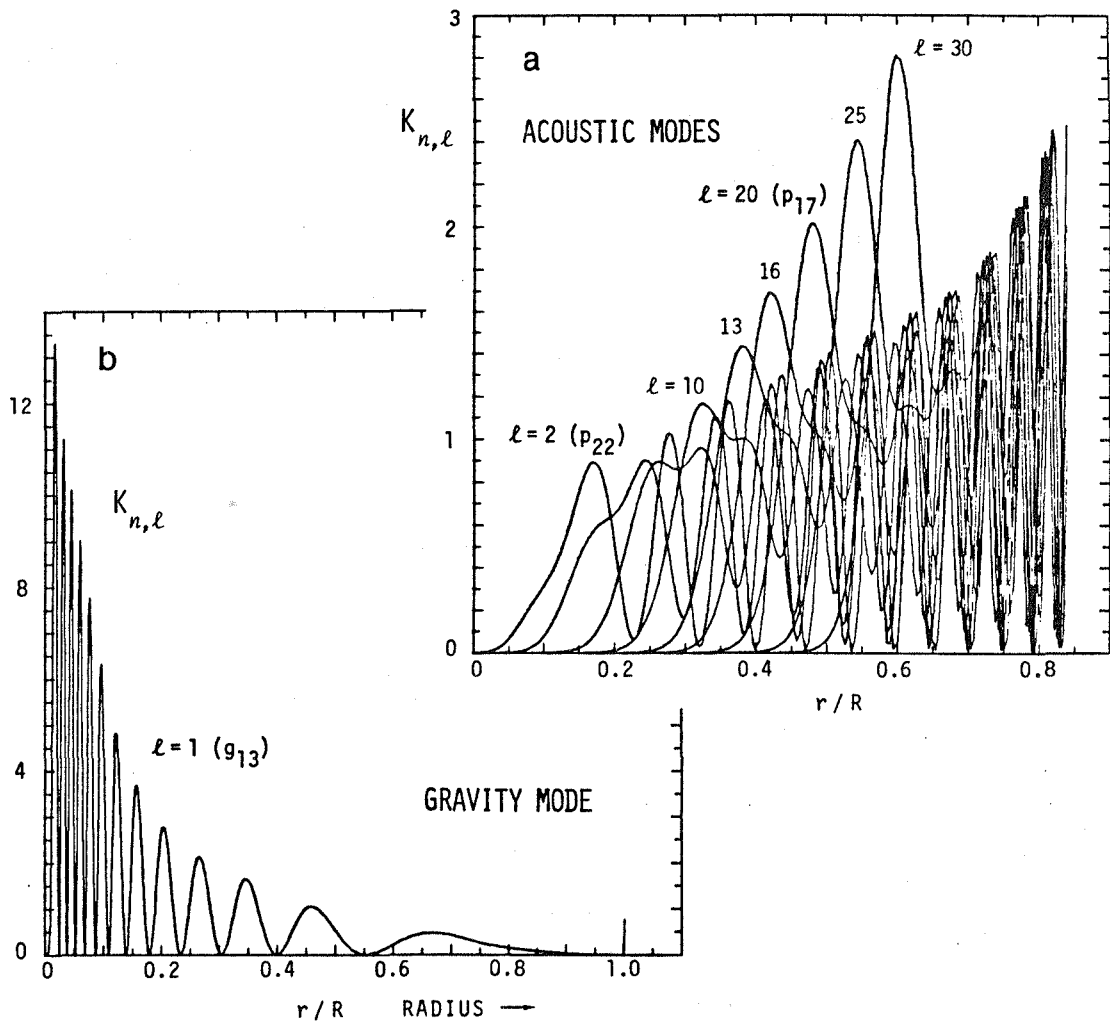
$$\nu_{n,\ell,m} = \nu_{n,\ell,o} + \frac{m}{2\pi} \bar{\Omega} , \quad (11)$$

where  $\bar{\Omega}$  is an average of the internal angular velocity of the form

$$\bar{\Omega} = \frac{\int_0^R \Omega(r) \frac{dr}{c}}{\int_0^R \frac{dr}{c}} . \quad (12)$$

If the Sun were to be rotating throughout at about its surface rate, the rotational splitting in frequencies between  $m$  components is about 0.46  $\mu\text{Hz}$ . That splitting is no longer simply a constant when one admits that  $\Omega$  is really a function of both  $r$  and  $\theta$ , and a few explicit examples have been considered (cf. Gough 1982a).

The rotational splitting kernels  $K_{n,\ell}$  shown in Figure 14 emphasize that a given  $p$  or  $g$  mode samples the angular velocity  $\Omega$  over a considerable range of depths. The  $g$  modes afford more information near the core and  $p$  modes in the upper regions, much as we surmised by



**Figure 14.** Variation with proportional radius  $r/R$  of rotational splitting kernels  $K_{n,l}$  deduced from theory, showing how selected acoustic and gravity modes might respond to an angular velocity  $\Omega(r)$  within the Sun. Shown in (a) are kernels for some p modes with periods close to 5 min, and in (b) for a g mode with period of 345 min. The kernels for these p modes become highly overlapped as they approach their maxima near the solar surface (attaining values close to 70 in the units shown); that region has been omitted here for clarity in displaying the kernels within the interior.

looking at the amplitude functions in Figure 9 from which such splitting kernels are derived. However, the kernels  $K_{n,\ell}$  are not at all like  $\delta$ -functions with radius. Therefore deductions about the variation of angular velocity within the Sun require the use of inversion techniques that combine information from many modes to obtain localized measures of  $\Omega$ . Such inversion methods to be applied to the data are discussed in detail by Gough (these proceedings) and are a critical element when interpreting the rotational splittings.

### c. Observed Splittings

Fine structure in the peaks has been detected in the observed power spectrum of low-degree p modes by Claverie et al. (1981), who have attributed this to rotational splitting of the frequencies. Claverie et al. further suggest that their measurements imply that a suitably weighted average of the interior angular velocity  $\Omega(r)$  of the Sun, much as  $\bar{\Omega}$  in (11), is about twice the value of  $\Omega$  at the surface. Gough (1982b) raises questions about this interpretation, since  $2\ell+1$  splitting components have been reported for a given  $\ell$  value, whereas only  $\ell+1$  of them should be detectable if the rotation axis of the interior coincides with that of the photosphere. However, if those axes do not coincide, then all the  $m$  components of a given  $\ell$  may be detectable, since issues of symmetry about the apparent equator at the surface no longer apply. In this case, however, one expects additional frequency splitting, so that in general more than  $2\ell+1$  components should be seen (Gough and Taylor 1984). The net result may appear as a modulation with time in the amplitudes of the fine structure peaks. Isaak (1982) has suggested that the additional components may be the consequence of an intense rotating magnetic core, such as has been postulated by Dicke (1979) to account for the 12.2 day periodic component in the Princeton oblateness data (e.g. Dicke and Goldenberg 1974, Dicke 1981). Gough (1982b) and Dicke (1982) have carried out analyses of the splittings that might be caused by an oblique magnetic solar rotator. Gough suggests that it is premature to conclude much about what causes the apparent frequency splittings before questions about the  $2\ell+1$  components in the spectrum are resolved by additional observations, such as those under way by Isaak and colleagues.

Measurements of apparent rotational splittings from observations of fluctuations in limb-darkening functions have been reported by Bos and Hill (1983) for a selection of tentatively identified low-degree p and g modes. From these splittings, Hill, Bos and Goode (1982) estimate the internal rotation and associated gravitational quadrupole moment  $J_2$  of the Sun, suggesting that general relativity may be in trouble. However, Gough (1982a) reports an alternative interpretation of those data using optimal averaging inverse methods and finds no fundamental conflict in the value of  $J_2$  with that supposed by Einstein in using general relativity to explain the advance of the perihelion of Mercury.

Certainly the subject of splitting of modal frequencies has engendered considerable debate, but that serves to emphasize just how

fundamental are the issues at stake, whether they touch on stellar structure, evolution or relativity. Clearly more observations, and preferably with a higher frequency resolution, are needed to try to resolve the interpretations or speculations put forward so far. Possibly detailed measurements of frequency splitting of the nonaxisymmetrical counterparts of the intermediate-degree modes, as in the observations briefly discussed by Harvey and Duvall (these proceedings; see also Duvall and Harvey 1984, Duvall et al. 1984) may help to sort out some of these issues.

## 5. PROBING FOR GIANT CONVECTION CELLS

The presence of large-scale convection cells below the solar surface should also lead to the splitting of frequencies, having particularly discernible effects on those modes whose horizontal scale is small compared to that of the cells. Magnetic field patterns have suggested the existence of such giant convection cells, as have nonlinear convection theories. Attempts to observe such convective flows by direct Doppler measurements have been inconclusive, and suggest that the photospheric velocity amplitude of the giant cells is below the sensitivity of the measurements, namely about  $10 \text{ ms}^{-1}$  (cf. LaBonte, Howard and Gilman 1983). However, theoretical modeling of compressible convection (e.g. Toomre 1980, Latour, Toomre and Zahn 1983) suggests that the horizontal velocity should increase with depth. Therefore one may anticipate finding velocities associated with the giant cells in the upper portions of the convection zone which are substantially greater than the limits set by the photospheric observations.

The detection and detailed measurement of large-scale convection cells would be essential for trying to unravel the complicated nonlinear dynamics of that zone. Most of what is seen at the surface of the Sun is controlled by underlying dynamical processes in the convection zone, whether by magnetic dynamo action that drives the solar cycle, or by convection under rotational constraints that redistributes the angular momentum and thus produces differential rotation, or by phase locking of convection cells that may explain the active longitudes and the persistent coronal holes, or by magnetoconvection that concentrates magnetic fields into flux ropes which erupt through the surface to form active regions. Although these are all formidable issues which will not be readily sorted out, seismology affords the possibility of observing convective flows and related structures below the surface where the dynamics is determined.

### a. Observation and Inversion of Data for Giant Cell Velocities

Acoustic modes of high degree  $l$  are most suitable for probing the giant cells, with the studies to date having used sectoral modes centered on the equator. The wave patterns are advected by horizontal flows which are composed both of rotation and convective cellular motions. The advection velocity  $\bar{U}$  for these high-degree modes is a spatial average of the equatorial component of the subphotospheric flow

velocity  $U$ , weighted by the energy density of the mode (cf. Gough 1978, Gough and Toomre 1983). If variations of  $U$  across the observing aperture and throughout each observing interval can be ignored, then the frequency of a given mode  $(n, \ell, m)$  is simply shifted by

$$\delta\nu = 2\pi k_h \bar{U} \quad , \quad (13)$$

and that shift is of opposite sign for the eastward ( $+k_h$ ) and westward ( $-k_h$ ) propagating modes. Temperature or density perturbations associated with such convection cells will produce further frequency shifts, but they should be independent of the sense of propagation of the modes. The frequency shifts attributable to giant cells are those that change over the course of different observing days as other convection cell patterns are rotated into view. Since individual modes of high degree cannot be resolved by the available observations, the frequency shifts  $\delta\nu$  are measured from the positioning of appropriate ridge sections in the  $k_h$ - $\nu$  power diagrams.

Hill, Toomre and November (1982, 1983) reported variations in ridge positions from one day to another that were thought to be associated with the passage of giant cells across the field of view. Figure 3 shows a portion of the power spectrum from one of their observing days. Application of an optimal averaging inversion procedure to that data reveals that  $\bar{U}$  changes by about  $100 \text{ ms}^{-1}$  from one observing day to another, and thus may be a consequence of giant cells (Hill, Gough and Toomre 1984). The available modes and the estimated signal-to-noise ratios permit inversion of the data to be carried out over a depth range of 15 Mm below the surface. Details of the sensitivity of such inversion methods are presented by Hill, Gough and Toomre (these proceedings). Although the use of high-degree  $p$  modes to study giant cells in the subphotosphere is still in preliminary stages, the method holds out very considerable promise.

#### **b. Influence of Atmospheric Seeing**

Observations of high-degree modes from the ground are impaired by the effects of atmospheric seeing. As discussed briefly by Hill (these proceedings), the image displacements and distortions arising from optical turbulence in the Earth's atmosphere lead to considerable degradation of the observed signal of 5-minute oscillations at the higher  $\ell$  values. This would explain the noticeable decrease of the power along the ridges with increasing  $k_h$  or  $\ell$ . Numerical simulations of seeing effects currently under way indicate that degradation sets in for  $\ell \gtrsim 200$ , and that power at  $\ell \approx 400$  is reduced by more than a factor of 20 even when atmospheric seeing conditions are good (Toomre et al. 1984). Seeing effects appear to introduce a background level of noise in the power spectra that may make the detection of low-amplitude modes at even low values of  $\ell$  difficult. The width and jaggedness of the ridges in Figure 3 are partly a consequence of seeing redistributing the power, and partly the result of beating between the unresolved modes within each sampling bin in  $k_h$  and  $\nu$ . Effects of mode beating can be reduced by data sets spanning longer intervals in time, though solar

rotation places limits of about 4 days over which the same pattern of giant cells might expect to be seen without too much distortion on the solar disk.

Networks of ground stations should remove effects of day-night gaps in such data, though successfully combining high  $l$  observations from different sites may prove to be difficult. However, all ground-based observations of high-degree modes are seriously affected by seeing distortions, and this is likely to place the main constraint on the effectiveness of solar seismology to study the dynamics of the convection zone. Such studies would benefit the most from observations carried out from space: a satellite positioned in a fully sunlit orbit would provide the highest signal-to-noise ratio in the measurement of high-degree oscillations while minimizing the observing interval necessary to accomplish this. It is also critical that such observations of dynamical structures be completed within their lifetimes, and this may preclude waiting for solar rotation to bring particular giant cells back into view again in order to continue the observations and thus try to beat down the noise in ground-based studies.

## 6. GROUND-BASED NETWORKS AND SPACE OBSERVATIONS

One can readily perceive that the study of solar oscillations by observation and theory has become an area of intensive and mature scientific inquiry. Progress in this subject has been striking. Solar seismology has already provided preliminary estimates for the initial solar helium abundance, has confirmed the essential features of temperature and density profiles within solar interior models, has determined the depth of the convection zone, has provided tentative evidence for possible mixing within the core, has yielded a preliminary determination of the rotation profile in the interior, and has provided evidence for subsurface flows associated with giant convection cells. Such a listing of the principal accomplishments of solar seismology to date is quite impressive. However, one should not assume that such an outpouring of results and interpretations will simply continue at an ever quickening pace. The reason for that note of caution is that progress in solar seismology will require increasingly precise determinations of oscillation frequencies, coupled with unambiguous identification of the modes. Neither ingredient is easy to achieve. Further, nonsolar noise must be at a low enough level that the frequencies of solar oscillations are distinct and unaltered. Also, the observations of dynamical structures within the Sun, such as of giant cells, must be accomplished before the structures have changed.

Several major steps are necessary for significant advances in solar seismology. One essential element will be the development of a new generation of very stable Doppler velocity analyzers that permit imaging of the Sun. The various studies under way on a Fourier tachometer operating as a Michelson interferometer, on a magneto-optical resonance line filter, on Fabry-Perot interferometers, on birefringent filters, and on stabilized grating spectrographs may provide several viable instruments. However, there is no paucity of challenges in getting such imaging



instruments to work with say  $0.5 \text{ ms}^{-1}$  absolute Doppler stability, for that implies 1 part in  $10^9$  stability in wavelength determination.

What are also needed are opportunities to carry out observations over long periods of time without significant interruptions due to day-night cycle or to weather. For instance, it would be essential to determine the frequencies of low-degree p modes to an accuracy of at least  $0.3 \text{ } \mu\text{Hz}$  in order to be able even to detect rotational splitting that is of the order  $0.46 \text{ } \mu\text{Hz}$  (if the Sun were to be simply rotating throughout at its surface rate). For typical frequencies of  $3000 \text{ } \mu\text{Hz}$  in the 5-minute band, this implies observations spanning at least 30 days to achieve the required frequency resolution, say  $\Delta\nu/\nu$ , of  $10^{-4}$  (recalling that  $\Delta\nu/\nu$  varies inversely with the number of oscillation periods observed coherently). Such frequency resolution probably is also necessary to sort out the very dense spectrum of gravity modes. For a g mode with a period of 160 min, or a frequency of about  $104 \text{ } \mu\text{Hz}$ , observations would be required that span at least 3 years. Yet clearly one would wish to have an even better frequency resolution to be able to determine how the fine structure varies with azimuthal order m, thereby being able to study differential rotation in both radius and latitude. Also, it would be essential to reduce the very complicated effects of day-night sidelobes (displaced from each real peak in the power spectra by  $\pm 11.6 \text{ } \mu\text{Hz}$ ) by having nearly continuous observations.

Finally, precise observations are required of high-degree modes if the dynamics of the convection zone are to be studied in any detail. Such observations would be essential in another respect: determination of conditions at considerable depths using the low-degree modes will be most reliable if accompanied by precise measurements of high-degree modes confined near the surface. Observations of the high-degree modes are influenced the most by seeing distortions introduced by viewing through a turbulent atmosphere, and significant progress on high  $l$  oscillations will have to await an observatory in space.

Progress in the long term in solar seismology will require developing both ground-based networks of observatories with imaging and refined space-based instruments. The networks of stations around the globe will permit nearly uninterrupted observing of the Sun, thereby largely removing effects of sidelobes in power spectra and considerably enhancing the signal-to-noise from what could be achieved from a single station in a comparable time span. Elimination of sidelobes will greatly aid the study of the splitting of frequencies by rotation and magnetic fields, and remove ambiguities in the frequency intervals populated by the closely-spaced g modes. The attainment of good signal-to-noise in reasonable observing intervals from networks of stations will permit the study of slowly time-varying stratification within the Sun during the course of the 11-year solar cycle. Ground stations are relatively easy to establish and they can be maintained over long intervals of time. Therefore it should be possible to begin the necessary observations quite soon. However, the primary emphasis in observations carried out from ground-based networks should be on the

measurement of oscillation modes of low and intermediate degree that are not seriously impaired by atmospheric seeing.

The space-based observatory will be essential for the accurate measurement of high-degree modes in order both to probe the dynamics of the convection zone and to remove ambiguities in the deep stratification as deduced from the low-degree modes. Without the high-degree modes, it will not be possible to analyze the crucial upper convective boundary layer at the top of the zone. Very accurate determination of frequencies for such modes should define the variation of sound speed with depth sufficiently well both to be able to specify the helium abundance of the convection zone and to calibrate equations of state of partially ionized plasma in the zone. Space observations will also permit the study of low-degree p and g modes of low radial order that are likely to be an important element in considering the stratification within the deep interior. Clearly observations from ground networks and from space are complementary, with the former easier to accomplish in the shorter term, but the latter essential for achieving the promise of seismology to probe a star in considerable detail. It is timely to proceed with both ventures.

#### ACKNOWLEDGMENTS

Extensive scientific advice and help was provided by Jorgen Christensen-Dalsgaard and Douglas Gough. Detailed discussions with Tim Brown, Tom Duvall, Jack Harvey, Frank Hill, George Isaak, John Leibacher, Bob Noyes, Ed Rhodes, Phil Scherrer, and Roger Ulrich were most instructive. This work was supported in part by the National Aeronautics and Space Administration through grants NSG-7511 and NAGW-91 and by the Air Force Geophysics Laboratory through contract F19628-82-K-0008.

#### FIGURE CREDITS

Fig. 1: Adapted from Musman and Rust (1970). Fig. 2: From Gough (1983), after Berthomieu et al. (1980), showing their theoretical frequencies superposed on observational power spectrum by Deubner (1975). Fig. 3: From Hill, Toomre and November (1983). Fig. 4: Adapted from Christensen-Dalsgaard (1982c). Figs. 5 and 6: From Gough (1983). Figs. 7, 8a, 9 and 14: Theoretical frequencies, eigenfunctions and rotational splitting kernels provided by J. Christensen-Dalsgaard and D. Gough, based on computations with a solar model (Model 1) described in Christensen-Dalsgaard (1982a). Fig. 8b: From Duvall and Harvey (1983). Fig. 10: From Christensen-Dalsgaard (1982b). Fig. 11: From Grec, Fosset and Pomerantz (1980). Fig. 12: Echelle diagram, plotted in the manner of Scherrer et al. (1983), showing theoretical frequencies computed from a solar model (Model 1) of Christensen-Dalsgaard (1982a) and observational South Pole data from Grec, Fosset and Pomerantz (1980). Fig. 13: Power spectrum courtesy of G. Isaak and colleagues.

## REFERENCES

- Ando, H., and Osaki, Y. 1977, *Publ. Astron. Soc. Japan* **29**, 221.
- Berthomieu, G., Cooper, A.J., Gough, D.O., Osaki, Y., Provost, J., Rocca, A. 1980, in **Nonradial and Nonlinear Stellar Pulsation: Lecture Notes in Physics 72**, (eds. H.A. Hill and W.A. Dziembowski; Springer), p. 307.
- Berthomieu, G., Provost, J., and Schatzman, E. 1984, *Nature* **308**, 254.
- Bos, R.J., and Hill, H.A. 1983, *Solar Phys.* **82**, 89.
- Brown, T.M., Mihalas, B.W., and Rhodes, E.J., Jr. 1984, in **Physics of the Sun** (eds. P.A. Sturrock, T.E. Holzer, D. Mihalas, and R.K. Ulrich; Reidel), in press.
- Christensen-Dalsgaard, J. 1982a, *Mon. Not. Roy. Astron. Soc.* **199**, 735.
- Christensen-Dalsgaard, J. 1982b, *Adv. Space Res.* **2**, 11.
- Christensen-Dalsgaard, J. 1982c, in **Proc. Pulsations in Classical and Cataclysmic Variable Stars** (eds. J.P. Cox and C.J. Hansen; JILA, Univ. Colorado), p. 99.
- Christensen-Dalsgaard, J., and Gough, D.O. 1976, *Nature* **259**, 89.
- Christensen-Dalsgaard, J., and Gough, D.O. 1981, *Astron. Astrophys.* **104**, 173.
- Christensen-Dalsgaard, J., and Gough, D.O. 1982, *Mon. Not. Roy. Astron. Soc.* **198**, 141.
- Claverie, A., Isaak, G.R., McLeod, C.P., van der Raay, H.B., and Roca Cortes, T. 1980, *Astron. Astrophys.* **91**, L9.
- Claverie, A., Isaak, G.R., McLeod, C.P., van der Raay, H.B., and Roca Cortes, T. 1981, *Nature* **293**, 443.
- Delache, P., and Scherrer, P.H. 1983, *Nature* **306**, 651.
- Deubner, F.L. 1975, *Astron. Astrophys.* **44**, 371.
- Deubner, F.L., and Gough, D.O. 1984, *Ann. Rev. Astron. Astrophys.* **22**, in press.
- Deubner, F.L., Ulrich, R.K., and Rhodes, E.J., Jr. 1979, *Astron. Astrophys.* **72**, 177.
- Dicke, R.H. 1979, *New Scient.* **83**, 12.
- Dicke, R.H. 1981, *Proc. Nat. Acad. Sci. USA* **78**, 1989.
- Dicke, R.H. 1982, *Nature* **300**, 693.
- Dicke, R.H., and Goldenberg, H.M. 1974, *Astrophys. J. Suppl.* **27**, 131.
- Duvall, T.L., Jr., Dziembowski, W.A., Goode, P.R., Gough, D.O., Harvey, J.W., and Leibacher, J.W. 1984, *Nature* **310**, 22.
- Duvall, T.L., Jr., and Harvey, J.W. 1983, *Nature* **302**, 24.
- Duvall, T.L., Jr., and Harvey, J.W. 1984, *Nature* **310**, 19.
- Fröhlich, C., and Delache, P. 1984, *Mem. Astron. Soc. Italiana* **55**, 99.

- Gabriel, M. 1984, *Astron. Astrophys.* **134**, 387.
- Goldreich, P., and Keeley, D.A. 1977, *Astrophys. J.* **212**, 243.
- Gough, D.O. 1978, in *Proc. Workshop on Solar Rotation* (eds. G. Belvedere and L. Paterno; Univ. Catania Press), p. 255.
- Gough, D.O. 1981, *Mon. Not. Roy. Astron. Soc.* **196**, 731.
- Gough, D.O. 1982a, *Nature* **298**, 334.
- Gough, D.O. 1982b, *Nature* **298**, 350.
- Gough, D.O. 1983, *Phys. Bull.* **34**, 502.
- Gough, D.O. 1984, in *Proc. Kunming Workshop on Solar Physics* (eds. B. Chen and C. de Jager), in press.
- Gough, D.O., and Taylor, P.P. 1984, *Mem. Astron. Soc. Italiana* **55**, 215.
- Gough, D.O., and Toomre, J. 1983, *Solar Phys.* **82**, 401.
- Grec, G., Fossat, E., and Pomerantz, M. 1980, *Nature* **288**, 541.
- Hill, F., Toomre, J., and November, L.J. 1982, in *Proc. Pulsations in Classical and Cataclysmic Variable Stars* (eds. J.P. Cox and C.J. Hansen; JILA, Univ. Colorado), p. 139.
- Hill, F., Toomre, J., and November, L.J. 1983, *Solar Phys.* **82**, 411.
- Hill, F., Gough, D.O., and Toomre, J. 1984, *Mem. Astron. Soc. Italiana* **55**, 153.
- Hill, H.A., Bos, R.J., and Goode, P.R. 1982, *Phys. Rev. Lett.* **49**, 1794.
- Isaak, G.R. 1982, *Nature* **296**, 130.
- Isaak, G., van der Raay, H.B., Palle, G., Roca Cortes, T., and Delache, P. 1984, *Mem. Soc. Astron. Italiana* **55**, 91.
- Kotov, V.A., Severny, A.B., and Tsap, T.T. 1978, *Mon. Not. Roy. Astron. Soc.* **183**, 61.
- LaBonte, B.J., Howard, R., and Gilman, P.A. 1983, *Astrophys. J.* **250**, 796.
- Latour, J., Toomre, J., and Zahn, J.-P. 1983, *Solar Phys.* **82**, 387.
- Leibacher, J.W., and Stein, R.F. 1971, *Astrophys. Lett.* **7**, 191.
- Leibacher, J.W., and Stein, R.F. 1981, in *The Sun as a Star* (ed. S. Jordan), NASA SP-450, p. 263.
- Leighton, R.B. 1960, in *Proc. IAU Symp. 12* (ed. R.N. Thomas; Suppl. Nuovo Cimento **22**), p. 321.
- Leighton, R.B., Noyes, R.W., and Simon, G.W. 1962, *Astrophys. J.* **135**, 474.
- Musman, S., and Rust, D. 1970, *Solar Phys.* **13**, 261.
- Rhodes, E.J., Jr., Ulrich, R.K., and Simon, G.W. 1977, *Astrophys. J.* **218**, 901.
- Scherrer, P.H. 1984, *Mem. Soc. Astron. Italiana* **55**, 83.

- Scherrer, P.H., and Wilcox, J.M. 1983, Solar Phys. **82**, 37.
- Scherrer, P.H., Wilcox, J.M., Christensen-Dalsgaard, J., and Gough, D.O. 1983, Solar Phys. **82**, 75.
- Scherrer, P.H., Wilcox, J.M., Kotov, V.A., Severny, A.B., and Tsap, T.T. 1979, Nature **277**, 635.
- Scuflaire, R., Gabriel, M., and Noels, A. 1982, Astron. Astrophys. **113**, 219.
- Severny, A.B., Kotov, V.A., and Tsap, T.T. 1976, Nature **259**, 87.
- Severny, A.B., Kotov, V.A., and Tsap, T.T. 1984, Nature **307**, 247.
- Tassoul, M. 1980, Astrophys. J. Suppl. **43**, 469.
- Toomre, J. 1980, Highlights Astron. **5**, 571.
- Toomre, J., Hill, F., Merryfield, W.J., and Gough, D.O. 1984, Solar Phys., submitted.
- Ulrich, R.K. 1970, Astrophys. J. **162**, 933.
- Ulrich, R.K., and Rhodes, E.J., Jr. 1983, Astrophys. J. **265**, 551.
- Vandakurov, Yu.V. 1967, Astron. Zh. **44**, 786.
- Willson, R.C., and Hudson, H.S. 1981, Astrophys. J. Lett. **244**, L185.
- Woodard, M., and Hudson, H.S. 1983, Nature **305**, 589.



## WHAT WOULD A DYNAMO THEORIST LIKE TO KNOW ABOUT THE DYNAMICS OF THE SOLAR CONVECTION ZONE?

Peter A. Gilman  
High Altitude Observatory  
National Center for Atmospheric Research

### 1. Current issues concerning solar dynamo theory.

Let me offer some fairly personal observations about the current status of solar dynamo theory. For recent detailed, balanced reviews, please see Stix (1981a), Schussler (1983), Weiss (1981), Cowling (1981), Belvedere (1983). I make no attempt here to be comprehensive, but rather shall try to highlight a few questions I think are of importance.

Until very recently, solar dynamo theory has been predominantly kinematic -- that is, the induction equation for magnetic field is solved using assumed velocities and/or parametric representations of the inductive or diffusive effects of velocities. The equations of motion governing these flows are not solved in parallel. Therefore nonlinear effects that have been invoked are generally *ad hoc*.

Within this kinematic domain, a number of apparent successes have been achieved, including magnetic field reversals of the correct period; migration of toroidal fields toward the equator to emulate the well known butterfly diagram; and others, including variations in cycle length and amplitude, when *ad hoc* nonlinearities are added (see e.g. Yoshimura 1978a, 1978b). But much of this apparent success results from the fact that there are free parameters and functions in the kinematic theory that can be chosen to yield the right results. In particular, to achieve the correct toroidal field migration direction requires that the angular velocity increase with depth; to get the correct dynamo period requires that the "helicity" (scalar product of velocity and vorticity) of the flow be weak, in some sense. But dynamical theories of the differential rotation driven by convection, at least those that take proper account of the influence of rotation upon convection, predict that the angular velocity should decrease with depth, and be nearly constant on cylinders concentric with the rotation axis. (See, e.g., Gilman, 1980 for a review.) And the helicity associated with the convection that drives the differential rotation is at least two orders of

magnitude larger than needed to give the correct period, so if field reversals occur at all, they are too fast by one order of magnitude (Gilman and Miller, 1981, Gilman, 1983).

Therefore there is currently a major paradox to be resolved in solar dynamo theory, one that has been known for almost a decade. Measurements of solar oscillations accurate enough to infer the interior rotation rate as a function of depth could largely resolve the paradox -- that is, whether kinematic dynamo theory is basically correct for the sun and the dynamical theories for global convection and differential rotation require major revision, or whether the dynamical theories are basically correct and solar dynamo theory is wrong.

If interior rotation can be measured from oscillations, then there is hope we can go further and attempt to demonstrate the existence of global convective flows, and hopefully infer some of their properties. Helicity would be an obvious candidate, though probably one of the most difficult.

If the dynamo theory proves to be wrong, it is probably either because it does not take into account in any detail the fact that the solar magnetic fields are highly concentrated into flux tubes, or because the dynamo is mainly concentrated in a thin layer at the bottom of the convection zone (e.g. Galloway and Weiss, 1981). It is not clear how oscillations measurements can help us understand these issues directly, since they can not give us highly localized information about a particular depth or structure in the convection zone.

## 2. Recent results from global compressible convection models.

Let me give a few more detailed results from the latest model calculations for global convection and differential rotation. These results are based on two models; my own, and that of Glatzmaier (1983). The physics in these models is rather similar, but the mathematical solution techniques are different. Both models are compressible, and typically contain several pressure scale heights. Both calculations begin significantly below the photosphere ( $\sim 10\%$  in solar radius) because they are incapable of resolving supergranule and granule scale convection. Both parameterize the effects of small scale (subgrid scale) eddy transports of momentum and entropy quite crudely. Among the physical differences between the models is the treatment of the lower boundary. My model has a rigid lower boundary, above which the stratification is all convectively unstable. Glatzmaier's model allows for penetration into a stable layer below. Glatzmaier's model also uses a spherical harmonic expansion in the horizontal dimension that allows better treatment of the poles than does mine, which uses a Fourier expansion in longitude, coupled with a finite difference grid in latitude and radius.

The differences between these models seem to be relatively unimportant in that the main results are quite similar. The fundamental conclusion from calculations using both these models of relevance to the sun is that, starting from



a solar luminosity, a solar rotation rate, and plausible, though crude, representation of the eddy diffusivities of momentum and entropy, they generate a broad spectrum of convection that in turn drives an equatorial acceleration of about the correct magnitude and profile. When one looks in a little more detail at the solution, some departures from solar observations do show up, but they do not negate the main agreement, suggesting we may be not too far from the "correct" answer for the sun (until one considers other issues, such as the dynamo problem as discussed above).

A number of additional results from these models should also be mentioned. From both, the angular velocity decreases with depth and is roughly constant on cylinders concentric with the axis of rotation. Convection in low latitudes is also predominantly in the form of convective rolls with a north-south axis. These results are both quite similar to that of the incompressible case studied earlier in a number of papers by me (see, e.g. Gilman 1979, 1980, and earlier references cited therein), and arise from the strong influence of rotation upon the flow.

In general, the spectrum of convection (measured, say, by the kinetic energy in each longitudinal wave number  $m$ ) is quite broad. In my calculations, the kinetic energy drops by a factor of 10 compared to the spectrum peak (somewhere in the range  $m = 10$  to  $15$ ), by  $m = 30$  to  $40$ . At certain times, however, there are exceptions, when a single longitudinal wave number contains a substantial fraction of the total convective kinetic energy. Glatzmaier found a case when wave number 14 contained 34% of the kinetic energy of the convection for a period of 2 months. No other wave number had more than 10% of the energy.

Within this broad spectrum, individual convective modes grow, evolve, and decay, and are advected by the differential rotation. From linear theoretical arguments (Glatzmaier and Gilman, 1981), they also are dispersive. That is, different wave numbers are predicted to rotate at different rates, by as much as several %. The result of all of these effects is that the total flow pattern changes radically in 1 rotation, and changes significantly in one passage from east to west limb ( $\sim 8$  days). We return to this point later.

Both helicity (defined above) and the Reynolds stress (covariance of perpendicular components of convective motion--a measure of the angular momentum transport by the convection) are prominent features of these solutions, with the helicity being a factor of  $\sim 10 - 10^2$  larger than required to give the correct dynamo period for the sun.

Two possible discrepancies with solar observations that show up are that the amplitude of the horizontal flow in the convection near the top of the model convection zone (but still  $\sim 10\%$  below the surface) is a factor of 2-3 larger than the upper limits estimated from the Mt. Wilson doppler velocity measurements by LaBonte *et al.* (1981). If both observations and theory are correct, then the global convection must be attenuated by the factor of 2-3 in the uppermost 10%

of the convection zone. Simple estimates of the amount of this attenuation to be expected (Stix, 1981b) are not encouraging.

Another difficulty is that we get better agreement in detail with the solar differential rotation if the convection zone is deep, extending, either as an unstable layer (my model), or through penetration (Glatzmaier model) to a depth of 40% or so of the radius. Shallower layers give a narrower equatorial acceleration, and some tendency to spin up the poles (my model). A convection zone depth or depth of convection mixing of 40% conflicts with estimates of burning of Li and Be (Vauclair, et al., 1978) and with observed five minute oscillation frequencies (see other papers in this proceedings volume). We continue to study solutions for shallower layers that do not have these conflicts, in search of better differential rotation profiles.

### 3. What we need to know about differential rotation and convection.

With respect to differential rotation, it is obvious from the above discussion that it is most important to find the profile of rotation with depth and latitude. If that is successful, then we can try for variations in rotation rate with time at the 1% level (20 m/sec). The recently discovered torsional oscillations (Howard and LaBonte, 1980) being at the 5 m/sec level, are almost certainly out of reach.

As everyone at this meeting knows, the first inferences of solar rotation with depth from frequency splitting of oscillations have given highly variable results. I suspect most workers in the field at this moment do not believe any of them very strongly. It seems imperative to make redundant calculations of rotation at the same depth using different combinations of oscillations, in order to test the reliability of the results. And we still need to demonstrate we can reproduce the known photospheric differential rotation with latitude by using high wave number p modes that peak near the solar surface.

With respect to convection, we would clearly like to know several properties. We would like to find the amplitude and structure with depth of the spectrum of convection out to longitudinal wave numbers of 30 or more. Does this structure yield north-south rolls in low latitudes as predicted? The uppermost 10% of the convection zone is of particular importance, because global models can not treat that layer well, and, as discussed above, significant attenuation with height is needed in convection amplitudes to fit with observations at the surface. To be useful, we need to measure convective velocities of the whole convection spectrum to an accuracy of 10-20 m/sec, particularly if the upper limit estimated from surface measurements by LaBonte *et. al.* (1981) is indicative of the interior flows. For a single longitudinal wave number, we would like to see velocities measured down to 5-10 m/sec, which may be very difficult.

Some properties of global convection expected on theoretical grounds may help us to infer their existence and flow structure from oscillations, while others may increase the difficulty of determining their nature. The favorable properties are particularly the expected symmetries of the various flow

components about the equator, while the unfavorable ones are the various sources of time dependence in the total convection pattern.

Theoretical models of global convection such as we have discussed above predict north-south rolls centered about the equator. This means that for a reasonably deep convection zone (30% of the solar radius) the east-west flow in the rolls at a given level should have the same sign and roughly the same amplitude from at least  $20^\circ$  N to  $20^\circ$  S. The radial motion will have the same symmetry. The north-south motion, on the other hand, will have a node at the equator, and have opposite signs at the same longitudes in the north and south. Thus if the convection velocities contribute a measurable splitting or shift to the oscillation frequencies, alternately adding and subtracting the signals in the north and south hemispheres ought to separate the contributions from the two horizontal velocity components. If this proves to be possible, then it will enhance our ability to demonstrate the existence of the north-south rolls, but also may allow calculation of higher order moments of the velocity field, such as the helicity, important for the dynamo, and the Reynolds stress, needed to understand how the differential rotation is maintained.

The longer the lifetime of a particular convection pattern (to which all longitudinal wave numbers contribute), the easier it is going to be to observe. Ideally, the same pattern recurring for several solar rotations could be superimposed to reduce noise. Even persistence during one passage from east to west limb would be valuable. However, there are at least three effects at work that limit pattern lifetime. These are a) individual mode growth and decay; b) advection and shearing of modes by the differential rotation, and c) dispersion of modes due to rotation. Let us take each of these in turn.

The lifetime of an individual convective mode in a turbulent system is closely related to its turnover time,  $L/U$ , in which  $L$  is a typical length scale for the convection, and  $U$  a typical velocity scale. If  $L = 0.1 R_\odot$  for large scale convection near the bottom of the convection zone, and  $U \sim 20$  m/sec, then  $L/U \sim 40$  days, or 1.5 solar rotation. By this measure, then, convection should easily persist through one passage from east to west limb (8 days) and there should be significant correlation from one rotation to the next.

Advection by the differential rotation at a given level should only change the speed of propagation of the convection pattern, and not appreciably affect its lifetime. However, shearing by differential rotation, if it is strong enough, can radically alter the structure of a convective mode. In Glatzmaier's and my global convection-differential rotation models both kinds of behavior can be seen. Latitudinal shearing is most important in mid latitudes, and weakest near the equator, so it will have the effect of limiting the latitudinal extent of persistent convection patterns. From my own model calculations, N-S rolls retain their shape well against latitudinal shearing equatorward of  $\sim 20^\circ$  latitude, but poorly poleward of that latitude. Between  $20^\circ$  and, say  $50^\circ$  latitude, there is a rotational velocity difference (relative to a uniformly rotating reference frame) of  $\sim 100$  m/sec. Two points, at  $20^\circ$  and  $50^\circ$  latitude initially on the same meridian, would therefore move  $30^\circ$  in longitude relative to each other in a time of

$\sim 0.5 R_{\odot} / 100 \text{ m/sec} = 40 \text{ days}$ . A convection pattern with a longitudinal wave number  $m = 12$  would be sheared by 1 full wave length in that time. Consequently a convective roll, representing half a convective wave length, would lose its structure in  $\sim 1/4$  that time, or 10 days. This time period would be shorter for higher wave numbers. Thus latitudinal shearing should prevent identification of a convective roll poleward of  $20^{\circ}$  latitude from one rotation to the next, and will greatly modify it in even a single disk passage ( $\sim 8$  days from 60 E to 60 W). This means that we should concentrate our effort on equatorial latitudes when looking for persistent convective features. As discussed above, it is also in low latitudes that most advantage can be taken of the predicted symmetry properties of the convection.

Linear theory of north south rolls in a highly stratified compressible fluid (Glatzmaier and Gilman, 1981) predicts there will be substantial dispersion in the longitudinal phase velocities of modes with different longitudinal wave number. It is not known yet how important these effects are in the nonlinear case, but from Glatzmaier and Gilman (1981) we can illustrate that this dispersion can lead to substantial change in a composite convection pattern in a single disk passage. They predict, for example, that a north-south roll with longitudinal wave number  $m = 12$  could have a period of as much as 6% less than  $m = 6$ . Thus in 8 days, while  $m = 6$  would have moved 2 of its wave lengths,  $m = 12$  would have moved only 3.69 of its, rather than the 4 wave lengths required to keep a composite pattern made up from these two modes invariant. Wavenumber  $m = 12$  lagging by 0.31 of its wave lengths means a phase lag of  $\sim 112^{\circ}$ , so the composite pattern must change substantially, if  $m = 6$  and  $m = 12$  have roughly comparable amplitudes. Nonlinear model results need to be analyzed in greater detail to see how important an effect this is. Finally, we note that additional dispersion will be introduced by the radial gradient of rotation if different longitudinal wave number rolls have peak amplitudes at different depths.

From all of the above discussions we conclude that composite convective patterns are unlikely to persist recognizably for even a single rotation, and may change significantly even during a single disk passage. However, if individual longitudinal wave numbers can be extracted from the overall pattern, these should last much longer. Changes in their amplitudes in low latitudes where shearing by latitude gradients in rotation is weak are governed by their turnover time previously estimated to be 1 - 2 rotations.

It would also be highly desirable to deduce from oscillations the horizontal and radial variations in thermodynamic structure caused by convection, but since these variations are likely to be of order of or less than  $10^{-5}$  everywhere except in the first few thousand km below the photosphere, the prospects do not seem good. Also, we would like to gain information on the magnetic field in the convection zone (and below) but it takes extremely large fields ( $\sim 10^6$  gauss) to produce a significant perturbation in the p mode frequencies at the base of the convection zone where strong fields might be able to reside. Thus, that prospect seems remote, too. But if the oscillations can be used to deduce the angular velocity profile with depth and latitude in the convection zone, that

will be a great advance indeed. And if even some of the velocity structure of convection can also be determined, progress will be much greater still.

#### ACKNOWLEDGEMENTS

I wish to thank Ronald Gilliland for reviewing the manuscript. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

## BIBLIOGRAPHY

- Belvedere, G. 1983, in *Proc. IAU Colloquium 71*, Eds. M. Rodono and P. Byrne, 579.
- Cowling, T.G. 1981, *Ann. Rev. Astron. Astrophys.* **19**, 115.
- Galloway, D.J. and Weiss, N.O. 1981, *Astrophys. J.* **243**, 945.
- Gilman, P.A. 1979, *Astrophys. J.* **231**, 284.
- Gilman, P.A. 1980, *Highlights of Astronomy* **5**, 91.
- Gilman, P.A. and Miller, J. 1981, *Astrophys. J. (Suppl.)* **46**, 211.
- Gilman, P.A. 1983, *Astrophys. J. (Suppl.)* Oct. issue (in press).
- Glatzmaier, G.A. 1983, *J. Computational Phys.* (submitted)
- Glatzmaier, G.A. and Gilman, P.A. 1981, *Astrophys. J. (Suppl.)* **45**, 381.
- Howard, R. and LaBonte, B.J., 1980, *Astrophys. J. (Lett.)* **239**, L33.
- LaBonte, B.J., Howard, R. and Gilman, P.A. 1981, *Astrophys. J.* **250**, 796.
- Schussler, M. 1983, in *Proc. IAU Symposium 102*, Ed. J. Stenflo, 213.
- Stix, M. 1981a, *Solar Phys.* **74**, 79.
- Stix, M. 1981b, *Astron. Astrophys.* **93**, 339.
- Vauclair, S., Vauclair, G., Schatzman, E., and Michand, G. 1978, *Astrophys. J.* **223**, 567.
- Weiss, N. 1981, in *Solar Phenomena in Stars and Stellar Systems*, Ed. R.M. Bonnet, 449.
- Yoshimura, H. 1978a, *Astrophys. J.* **220**, 692.
- Yoshimura, H. 1978b, *Astrophys. J.* **226**, 706.

## SOLAR INVERSE THEORY

Douglas Gough

Institute of Astronomy and Department of Applied Mathematics and Theoretical Physics, University of Cambridge, and Joint Institute for Laboratory Astrophysics, University of Colorado and National Bureau of Standards, Boulder, Colorado 80309

Helioseismological inversion, as with the inversion of any other data, is divided into three phases. The first is the solution of the so-called forward problem: namely, the calculation of the eigenfrequencies of a theoretical equilibrium state. The second is an attempt to understand the results, either empirically by determining how those frequencies vary as chosen parameters defining the equilibrium model are varied, or analytically from asymptotic expansions in limiting cases of high order or degree. A familiarity with at least the qualitative dependence of the eigenfrequencies on various properties of the solar model is necessary not only for personal enlightenment but also for arming oneself to interpret the rather more abstract third phase. That phase is to pose and solve an inverse problem, which seeks to find a plausible equilibrium model of the Sun whose eigenfrequencies are consistent with observation.

The three phases are briefly discussed in this review, and the third, which is not yet widely used in helioseismology, is illustrated with some selected inversions of artificial solar data.

### 1. INTRODUCTION

There is now a substantial body of seismological data from the Sun. Therefore we can contemplate graduating from the hit-and-miss model-fitting investigations that have dominated the subject in the past to the more systematic approach provided by inverse theory. There are many methods for attempting to invert seismological data, developed mainly by geophysicists and applied mathematicians interested in geophysical problems, so we are in the fortunate situation of being able to draw on their experience. There appears to be a variety of schools of opinion about which methods are generally preferable, or even which approach is likely to be the most fruitful for a particular problem. What seems to be universally agreed, however, is that when attacking a totally new problem it is expedient to test a method first, using artificial data generated from a theoretical model.

In this and the accompanying discussion (Christensen-Dalsgaard and Gough, 1984a), we report briefly on two potentially useful methods of inverting helioseismological data. One is based on localized averages, and the other is an expansion (spectral expansion) in the weighting functions (data kernels) that characterize the contribution of different regions of the star to the data. Both methods have been used in geophysics, and have their origins in the work of Backus and Gilbert (e.g. 1970). A detailed description of the methods, and their relation to other techniques that have been used, will not be given here; for this the reader is referred to the various reviews that already exist (e.g. Wiggins, 1972; Parker, 1977a,b; Sabatier 1977, 1978). The examples discussed here and in the accompanying article of the application of these methods to artificial data suggest that both of the procedures offer hope of being useful for inverting solar data.

The two aspects of the solar structure that are most likely to be determined in the foreseeable future are the internal rotation and the hydrostatic stratification. From a theoretical point of view the former is the simpler problem, because it is essentially linear; moreover the result is likely to be quite reliable, because it is not sensitive to uncertainties in the physics of stellar material. The latter is non-linear, and depends on a knowledge of the equation of state and, to a lesser extent, on the dynamics of convection. Nevertheless there are good grounds to believe that a useful inversion, which sets certain stringent constraints on the stratification, will soon be possible.

## 2. PHASE ONE: THE FORWARD PROBLEM

The forward problem is discussed in the standard texts on stellar pulsation theory (Ledoux and Walraven, 1958; Unno *et al.*, 1979; Cox, 1980). Since to a first approximation the Sun may be regarded as being spherically symmetrical (the oblateness of the density distribution is nowhere more than about  $10^{-5}$ ), the oscillation eigenfunctions are separable. For example, the displacement eigenvector  $\xi(\mathbf{r}, t)$  may be written, with respect to spherical polar coordinates  $(r, \theta, \phi)$ :

$$\xi = \left( \Xi(r) P_{\ell}^m, r^{-1} H(r) \partial_{\theta} P_{\ell}^m, r^{-1} \operatorname{cosec} \theta H(r) P_{\ell}^m \partial_{\phi} \right) \cos(m\phi - \omega t) \quad (2.1)$$

where  $P_{\ell}^m(\cos\theta)$  is the associated Legendre function of the first kind and  $t$  is time;  $\ell$  and  $m$  are respectively the degree and the azimuthal order of the mode.

To a good first degree of approximation (no worse than 1 part in  $10^3$ ) the oscillations may be regarded as being adiabatic. This property has a good aspect and a bad aspect. The good aspect is that the physics of adiabatic oscillations is relatively well understood. So an accurate and fairly reliable solution to the forward problem is possible. The dynamics of the oscillations hardly senses the perturbations to the heat flow, and therefore it is not necessary to have an accurate knowledge of the opacity or the nuclear energy generation rates, or how the convective



heat flux is modulated by the oscillations, at least for the lower-frequency modes. The highest frequencies that have been observed do seem to have a somewhat greater sensitivity to the nonadiabatic effects of convection, however, the theory of which is quite uncertain. Nevertheless, since these are significant only in a very thin layer near the top of the convection zone, it is possible, at least in principle, to eliminate their influence by considering appropriate combinations of eigenfrequencies. I shall not discuss that here, but merely assume that they can be neglected, and consider the straightforward linearized perturbation analysis of a truly hydrostatic star.

The bad aspect relates to the good aspect: because heat flow is essentially irrelevant, the oscillation eigenfrequencies provide us with no direct information about temperature. Adiabatic oscillations result solely from Newton's second law of motion: pressure gradients exert a force on material with inertia. Therefore all we can learn directly is a relation between pressure and density. If this is known throughout the Sun it can be coupled with the constraint of hydrostatic balance and Newtonian gravitation to provide us, at least in principle, with the density and pressure stratification. Of course it is necessary to know the equation of state, so that one can calculate the perturbation to the pressure gradient associated with a given compression or rarefaction, and this is perhaps the major uncertainty in the entire theory at present.

The dominant symmetry-breaking agent is the linear advection term coming from the angular velocity of the Sun. Its effect is to split the degeneracy of the eigenfrequencies of like order ( $n$ ) and degree ( $l$ ). From the splitting one can hope to infer the internal angular velocity  $\Omega(r, \theta, t)$ . Because the splitting is small, quadratic effects, such as centrifugal forces acting on both the equilibrium state and the oscillations, can be ignored. Therefore the problem is, to a good approximation, linear. Several reports of observations of rotational splitting have already been made (Deubner *et al.*, 1979; Claverie *et al.*, 1981; Hill *et al.*, 1982; Deubner, 1983; Rhodes *et al.*, 1983a; Scherrer and Delache, 1984), and more are no doubt imminent.

In addition to  $\Omega$ , large-scale convection currents (giant cells) should also produce diagnostically useful observable consequences (Gough and Toomre, 1983; Hill *et al.*, 1983, 1984). Locally, the horizontal component of the convective flow acts on high-degree oscillations in much the same way as rotation.

### 3. PHASE TWO: UNDERSTANDING THE FORWARD PROBLEM

Much of the understanding of the forward problem comes from analytical analyses in simplified circumstances. Useful information about high-degree oscillations, for example, can be extracted by approximating the equilibrium state by a polytrope or by assuming it to be isothermal (e.g. Lamb, 1932; Stein and Leibacher, 1974). Low-degree oscillations can be analyzed asymptotically at high order (e.g. Vandakurov, 1967; Zahn, 1970; Ledoux and Perdang, 1980; Tassoul, 1980). It is from analyses such as

these that one can learn what aspects of the observed dispersion relation pertain to different levels in the star. In particular, the small separation between the frequencies of modes  $(n, \ell)$  and  $(n-1, \ell+2)$  of low-degree five-minute oscillations provides information about the stratification in the energy-generating core (e.g. Gough, 1983a). Another example is presented in these proceedings, where it is pointed out that the latest observations of five-minute oscillations by Harvey and Duvall (1984) indicate an error in a standard solar model at or near the base of the convection zone (Christensen-Dalsgaard and Gough, 1984b).

Asymptotic relations usefully provide a coarse diagnostic of the solar interior. The frequencies of high-order modes depend predominantly on aspects of the equilibrium structure that vary on a scale much greater than the characteristic wavelength of the oscillation eigenfunctions, and can be expressed as integrals of the equilibrium structure that do not depend explicitly on the detailed variation of these eigenfunctions. In particular, Duvall's law (1982) describes the p modes, and appears to be approximately true for high-degree modes even when the order is not large (Christensen-Dalsgaard *et al.*, 1984). The law can be inverted analytically to obtain an integral formula for the sound speed in terms of the observed dispersion relation (Gough, 1984).

Aside from discontinuities, aspects of the equilibrium structure that vary on a scale much shorter than a wavelength are not easily treated by analytical methods. They do influence the eigenfrequencies, however. In such cases, the effect on the frequencies depends on the detailed structure of the spatial oscillations in the eigenfunctions. Though it is often possible to see a posteriori how the frequencies are influenced, by noticing the locations of nodes and antinodes of the eigenfunctions in relation to the small scale variations in the equilibrium structure, it is usually not easy to make quantitative predictions.

It is certainly evident, therefore, that a systematic procedure is required for analyzing the eigenfrequencies to measure the equilibrium structure on a relatively small scale. In the present state of development of the theory, it is probably necessary to use such a procedure for much of the large-scale structure too, since the asymptotic approximations are not always reliable. Inverse theory will no doubt provide that procedure. But before embarking on my discussion of inverse theory it is perhaps appropriate simply to mention some of the conjectures that have arisen from the results of model-fitting. I do this to emphasize the importance of a simple procedure that provides an easy guide to the results that might subsequently be found by more sophisticated analyses.

The first conjecture from model-fitting was deduced from the fact that the frequencies of high-degree five-minute oscillations observed by Deubner (1975) were lower than the predictions (Ulrich, 1970; Ando and Osaki, 1976) of contemporary solar model envelopes. From a polytropic analysis (Gough, 1977) this seemed to imply that the models overestimated the entropy jump across the upper superadiabatic convective boundary layer, implying that the Sun's convection zone was actually deeper than the models predicted. Numerical experiments by Ulrich and

Rhodes (1977) were consistent with this idea, and subsequent more detailed numerical work (Lubow et al., 1980; Berthomieu et al., 1980), in which many of the other uncertain aspects of the theory were varied, suggested that the conjecture was likely to be the only possibility.

A second conjecture arose from a calibration performed by fitting to observations the eigenfrequencies of low-degree five-minute modes of a sequence of evolved solar models with different compositions (Christensen-Dalsgaard and Gough, 1980, 1981; Shibahashi et al., 1983; Ulrich and Rhodes, 1983). This suggested an initial solar helium abundance of  $0.25 \pm 0.02$  (see also Gough, 1983b). Interestingly, it was found that a perfect fit between theory and observation was not possible, and subsequent numerical experiments with the theory failed to remove the discrepancy. As I have already pointed out, we have a strong indication of where the models are incorrect (Christensen-Dalsgaard and Gough, 1984b), but the nature of the error is yet to be found. Perhaps we must await an inverse calculation before we know the answer.

#### 4. WHAT DO OSCILLATION FREQUENCIES TELL US?

Oscillations sample an extended region of the Sun, and so provide an integral measure of the structure over that region. This is most easily seen from the variational formulation of the adiabatic eigenvalue problem for a nonrotating star (e.g. Ledoux and Walraven, 1958):

$$\omega^2 \int \rho \xi \cdot \xi dV = \int [\gamma p (\text{div} \xi)^2 + 2 \xi \cdot \nabla p \text{div} \xi + \rho^{-1} \xi \cdot \nabla p \xi \cdot \nabla \rho] dV - G \iint |\underline{r} - \underline{r}'|^{-1} \text{div}(\rho \xi) \text{div}'[\rho(\underline{r}') \xi(\underline{r}')] dV' dV, \quad (4.1)$$

where  $p$  and  $\rho$  are the pressure and density of the equilibrium state,  $\gamma$  is the adiabatic exponent  $(\partial \ln p / \partial \ln \rho)_{ad}$ ,  $G$  is the gravitational constant and the integrals are over the volume of the star. Thus the frequencies are a combination of weighted averages of nonlinear functions of the equilibrium pressure and density and their derivatives, the weighting depending on the oscillation displacement eigenfunctions  $\xi$ . Since different eigenfunctions weight the structure differently, the hope is that with a sufficient variety of data one can obtain an estimate of how  $p$  and  $\rho$  vary with  $r$ .

Most of the inversion procedures that have been developed apply only to linear integral equations. Therefore to make some progress it is expedient to develop an iterative procedure, like a generalized Newton-Raphson method, that improves (hopefully) upon a trial model of the Sun. One postulates an initial guess:  $p_0, \rho_0, \gamma_0$ , taken, for example, from a standard model of the Sun. Then one carries out the forward problem, calculating the eigenfunctions  $\xi_0$  and eigenfrequencies  $\omega_0$  corresponding to the modes for which observational data are available. One presumes that the physics is correctly described by equation (4.1), so that  $\omega, p, \rho, \xi$  refer to the actual Sun. Then one writes down the equation satisfied by  $\delta \omega^2 \equiv \omega^2 - \omega_0^2$ , by subtracting two equations of the type (4.1),

and hopes that the trial model is sufficiently close to reality for linearization in the differences  $\delta p \equiv p - p_0$ ,  $\delta \rho \equiv \rho - \rho_0$ , etc. to be valid. There results an equation in which  $\delta \omega^2$  is expressed as a linear functional of the differences  $\delta p$ ,  $\delta \rho$ , etc. It is at this point that one appreciates the variational formulation of the problem. Since equation (4.1) is stationary to variations in the functions  $\xi$  about the true eigenfunctions, it follows that the terms that are linear in  $\xi - \xi_0$  cancel, and  $\delta \omega^2$  can be expressed in terms of  $\xi_0$  alone. Thus it is not necessary to perturb the forward problem.

One can proceed further by imposing the constraint of hydrostatic support

$$\frac{dp}{dr} = - \frac{G\tilde{m}\rho}{r^2} \quad , \quad (4.2)$$

where  $\tilde{m}$  is the usual mass variable which satisfies

$$\frac{d\tilde{m}}{dr} = 4\pi r^2 \rho \quad . \quad (4.3)$$

By substituting these equations into the equation for  $\delta \omega^2$  and integrating by parts (and noticing that the surface integrals are negligible), it is possible to write the equation in the form

$$\frac{\delta \omega^2}{\omega_0^2} = \int S(\underline{x}_0, \xi_0, r) \frac{\delta \rho}{\rho_0} r^2 \rho_0 dr \quad , \quad (4.4)$$

provided the equation of state is known, where  $\underline{x}_0(r)$  represents the equilibrium structure ( $p_0$ ,  $\rho_0$ , etc.) of the trial model. The formula for the differential kernel  $S$  is quite complicated, so I refrain from presenting it here. Several examples are plotted by Gough (1978a).

If equation (4.4) can be inverted to estimate  $\delta \rho$ , an improved estimate of  $\rho$  can be deduced, and the whole procedure could be repeated.

In deriving equation (4.4) I presumed the equation of state was known. This is required for computing the variation in  $\gamma$ . What is required is not only a knowledge of the microphysics of the solar material, but also the composition. Of course for the first iteration one has a trial composition taken from the standard solar model. But for subsequent iterations one has no such information, and the procedure must be generalized.

For the purposes of inferring the structure of the core (if it can be done without a detailed knowledge of the structure of the envelope) this issue is not very important. There the material is highly ionized, and  $\gamma \approx 5/3$ . But in the convection zone where the abundant elements are partially ionized, the problem is important. Formally it is quite straightforward to overcome the difficulty by generalizing equation (4.4) to include additional integrals that are weighted averages of the abundances of the elements that influence  $\gamma$ , and generalizing the inversion procedure described in the following section. [The geophysical

inverse problem originally formulated by Backus and Gilbert (1967) was posed for a vector function.] No investigation of the properties of such a procedure applied to the Sun has yet been undertaken. Alternatively, it may be simpler to separate the problem of determining the composition of the convection zone and solving that first (Gough, 1984). This should be possible because in the convection zone the stratification is close to being adiabatic and the chemical composition is presumably homogeneous.

The case of rotational splitting of nonaxisymmetrical modes can be treated similarly. The perturbations  $\omega_1$  to the eigenfrequencies imposed by the angular velocity  $\Omega(r, \theta, t)$  can be computed by linearizing the variational principle of Lynden-Bell and Ostriker (1967). Alternatively it can be obtained in the form of a consistency condition if one expands the eigenvalue problem about that for the nonrotating state. The result is a functional of  $\Omega$  which, in the special case when  $\Omega$  is a function of  $r$  alone, can be written in the form (Gough, 1981; cf. Hansen *et al.*, 1977)

$$\frac{\omega_1}{\omega_0} = m \int K(\chi_0, \xi_0, r) \Omega(r) dr \quad (4.5)$$

Quadratic and higher order terms can be computed, if desired, but these are small compared with the linear term represented in equation (4.5) (Dziembowski and Goode, 1984; Gough and Taylor, 1984). Once again, the data, namely the splitting frequencies  $\omega_1$ , are weighted averages of the function to be determined (this time  $\Omega$ ), with weight functions that depend on the equilibrium model and its eigenfunctions. If the previous inverse problem has already been solved (using the observed frequencies of the axisymmetrical modes)  $K$  can be regarded as being known. This inverse problem is therefore linear, and requires no iteration.

## 5. PHASE THREE: INVERSE METHODS

In this section I consider explicitly the inversion of the idealized equation (4.5) to obtain  $\Omega$  from rotational splitting data. Now it is convenient to label the modes with an index  $i$  which identifies  $n$ ,  $\ell$  and  $m$ . Actually, in the special case considered here where  $\Omega$  depends only on  $r$ ,  $K$  is independent of  $m$  and therefore  $i$  need represent only  $n$  and  $\ell$ . Thus one writes

$$w_i = \int_0^R K_i(r) \Omega(r) dr \quad (5.1)$$

where  $w_i = \omega_{1i}/(m\omega_{0i})$  are observable quantities,  $K_i$  are known kernels and  $R$  is the radius of the Sun. In practice, observations contain errors, so if the  $w_i$  are regarded as the observations, equation (5.1) holds only approximately.

As pointed out by Backus and Gilbert (1967), a solution to the problem represented by (5.1), if it exists, is not unique, even if strict equality is assumed. One reason is that since one can have only a finite number of observations, the kernels  $K_i$  cannot possibly span the space of functions on the interval  $(0,R)$ . Therefore there is an infinite number of functions,  $f_k$ , orthogonal to all the  $K_i$ , and any linear combination of them can be added to  $\Omega$  without modifying any of the integrals. Let  $O$  denote the subspace spanned by the  $K_i$ . The functions  $f_k$  lie in its complement,  $A$ , called the annihilator. Evidently the data contain no information about the projection of  $\Omega$  in  $A$ .

Permitting only approximate equality broadens the possibilities further. Therefore the problem is not simply to find an approximate solution to equation (5.1), but to select which of the infinite number (if there are any at all) is the most likely. Here prejudice reigns, and opinions therefore differ.

Before proceeding I shall assume that the physics embodied in equation (5.1) is correct (and that the data are consistent). In that case a solution must exist.

### Spectral expansion

This is essentially an expansion of  $\Omega$  in terms of the kernels  $K_i$ . The idea of using  $K_i$  as a basis is, at first sight, natural, since the  $K_i$  span the subspace  $O$  that is accessible to the observations. However, as I shall soon discuss, it is actually more useful to transform, at least conceptually, to a new basis,  $\psi_i$ , that takes into account the degree to which the data can measure the projection of  $\Omega$  onto each basis function.

In geophysics, the expansion first arose out of a procedure formulated by Backus and Gilbert (1967). Suppose one has a preconceived idea,  $W(r)$  say, of the function  $\Omega$ . Let us assume that  $W$  is close to the truth, and seek that function  $\Omega$  that minimizes the least-squares deviation

$$E \equiv \int_0^R (\Omega - W)^2 dr \quad . \quad (5.2)$$

Forgetting errors for the moment, the minimization must be performed subject to the constraints (5.1), with exact equality. The result is

$$\Omega = W + \sum_i \alpha'_i K_i \quad , \quad (5.3)$$

with  $\alpha'_i$  being the solutions of the Euler equations

$$\sum_j A_{ij} \alpha'_j = w_i - a_i \quad , \quad (5.4)$$

where

$$A_{ij} = \int_0^R K_i K_j dr \quad , \quad a_i = \int_0^R K_i W dr \quad . \quad (5.5)$$

Thus the difference between  $\Omega$  and  $W$  is expressed, in equation (5.3), as a linear combination of the kernels  $K_i$ .

The foregoing analysis suggests that one might simply express  $\Omega$  directly as an expansion in  $K_i$ :

$$\Omega = \sum_i \alpha_i K_i \quad . \quad (5.6)$$

If once again errors are ignored, one can determine the coefficients in the usual way by projecting the constraints (5.1) onto the basis, which leads to

$$\sum_j A_{ij} \alpha_j = w_i \quad . \quad (5.7)$$

The difference between the solutions (5.6)-(5.7) and (5.3)-(5.4) is in the annihilator  $A$ , so the data  $w_i$  cannot distinguish between the two. Therefore, if  $w_i$  constitute the only information that is available about  $\Omega$ , one has no basis for choosing between the two possibilities. An inherent advantage of the Backus-Gilbert formulation, however, is that one can write into  $\Omega$ , via  $W$ , additional constraints (obtained from information other than  $w_i$ ) that are not otherwise easily incorporated into a procedure for solving (5.1).

There are two related problems that one would encounter in trying to carry out the straightforward method outlined above. First, if there are redundant data, the matrix  $A_{ij}$  is singular, and some care must be taken in solving equations (5.4) or (5.7). If the data  $w_i$  were truly error-free, that would be possible, at least in principle, since, by hypothesis, the equations would be consistent, and one would need only to reject the redundant equations and solve the reduced set that remains. In practice, however, the data are erroneous, and equations (5.4) or (5.7) are formally inconsistent. One could still reject redundant data, but that, of course, would be unwise. Retention of redundancy is always important under these circumstances for reducing the influence of random errors. Therefore some kind of averaging procedure is required.

The second problem concerns error magnification. In addition to formal redundancy it is usually the case that there are different combinations of the data that give nearly but not strictly the same information. This leads to the matrix  $A_{ij}$  being ill-conditioned, or nearly singular. Once again, one wishes to average in some way the almost identical information that is contained in the two or more combinations; that is analogous to what must be done for the genuinely redundant data. But in addition one must reject the apparent information contained in the difference between the almost equivalent combinations, for that is dominated by the errors in the data.

The second problem is analogous to trying to measure a vector  $\underline{y}$  by making independent measurements of its components  $(\alpha_1, \alpha_2)$  in the directions of the unit vectors  $\underline{y}_1$  and  $\underline{y}_2$ , which are known to be nearly parallel. Thus if

$$\tilde{v} = \alpha_1 \tilde{v}_1 + \alpha_2 \tilde{v}_2 = \frac{1}{2} (\alpha_1 + \alpha_2) (\tilde{v}_1 + \tilde{v}_2) + \frac{1}{2} (\alpha_1 - \alpha_2) (\tilde{v}_1 - \tilde{v}_2) \quad (5.8)$$

and the measurements of  $\alpha_1$  and  $\alpha_2$  have errors  $\epsilon_1$  and  $\epsilon_2$ , one cannot measure the component of  $\tilde{v}$  in a direction roughly perpendicular to  $\tilde{v}_1$  and  $\tilde{v}_2$  (e.g. in the direction of  $\tilde{v}_1 - \tilde{v}_2$ ) if either of  $\epsilon_1$  or  $\epsilon_2$  has a magnitude comparable with the difference  $\alpha_1 - \alpha_2$  between the two measurements. Even though formally the vectors  $\tilde{v}_1$  and  $\tilde{v}_2$  can be used as basis vectors for a plane, in practice the erroneous data provide information only along the line parallel to  $\tilde{u}_1 = (1/2)(\tilde{v}_1 + \tilde{v}_2)$ . For practical purposes the component of  $\tilde{v}$  parallel to  $\tilde{u}_2 = (1/2)(\tilde{v}_1 - \tilde{v}_2)$  is inaccessible. Thus the situation is essentially the same as if  $\tilde{v}_1$  and  $\tilde{v}_2$  were genuinely parallel, which is analogous to the first of the problems mentioned above.

The resolution of problems of this kind had already been discussed by Lanczos (1961). The procedure is to find that part of the subspace  $O$  that is inaccessible to the data by virtue of the errors, and to relegate it to the annihilator  $A$ . The solution  $\Omega$  is the component of the actual function that is in the appropriately diminished subspace  $O'$  that remains. In terms of the simple vector analogy, one creates a new orthogonal basis  $(\tilde{u}_1, \tilde{u}_2)$  of the plane, and recognizes that one can measure only the component  $\alpha_1 + \alpha_2$  in the direction of  $\tilde{u}_1$ . Notice that in this analogy the result contains both measurements, so statistical errors are lower than had the second of the measurements, say, been rejected as being redundant and the resulting component estimated by  $2\alpha_1$ . The Lanczos inverse also has this property, providing a natural way of averaging data in the more complicated case when several different combinations provide similar information.

For the details and the justification of the Lanczos method the reader is referred to the book by Lanczos (1961), and the discussions in the geophysical context by Jackson (1972), Wiggins (1972) and Parker (1977a). The following summary follows Parker (1977a).

Let  $\sigma_i$  be the standard errors of the data  $w_i$ . Then one weights the constraints (5.1) with  $\sigma_i^{-1}$ , yielding

$$w_i' = \int_0^R K_i'(r) \Omega(r) dr \quad , \quad (5.9)$$

where  $w_i' = w_i/\sigma_i$  and  $K_i' = K_i/\sigma_i$ , so that  $w_i'$  has unit standard error. The matrix  $A_{ij}'$ , defined in the same way as  $A_{ij}$  in equations (5.5) but with  $K_k'$  replacing  $K_k$ , is positive-definite and symmetric, and can be diagonalized with an orthonormal matrix  $U_{ij}$  to give

$$\sum_{k,\ell} U_{ki} A_{k\ell}' U_{\ell j} = \lambda_i \delta_{ij} \quad (5.10)$$

where  $\lambda_i$  are the positive eigenvalues of  $A_{ij}'$  and  $\delta_{ij}$  is the Kronecker delta. Now consider the new basis  $\psi_i(r)$  of  $O$  defined by



$$\psi_i = \lambda_i^{-1/2} \sum_j U_{ji} K_j' \quad , \quad (5.11)$$

which has the property

$$\int_0^R \psi_i \psi_j dr = \delta_{ij} \quad . \quad (5.12)$$

Regard the functions  $\psi_i$  to be ordered such that  $\lambda_i$  decreases with increasing  $i$ , and expand the solution  $\Omega$  in terms of them:

$$\Omega = \sum_i \alpha_i \psi_i \quad . \quad (5.13)$$

Of course the expansion excludes that part of  $\Omega$  in the annihilator  $A$ . In view of equations (5.12) and (5.9), the expansion coefficients are determined by

$$\lambda_i^{1/2} \alpha_i = \lambda_i^{1/2} \int_0^R \psi_i \Omega dr = \sum_j U_{ji} w_j' \quad . \quad (5.14)$$

Recalling that the  $w_j'$  have unit standard errors, it follows immediately from equation (5.14) and the orthonormality of  $U_{ij}$  that if the errors in  $w_j'$  are statistically independent the errors in  $\alpha_i$  are also statistically independent, with standard deviation  $\lambda_i^{-1/2}$ . Thus the uncertainty of the coefficients  $\alpha_i$  increases with decreasing  $\lambda_i$ ; and it is immediately evident from equation (5.14) that the uncertainty is total when  $\lambda_i = 0$ . Notice that these statements incorporate the statistics of the errors in the data, as must be the case, because each  $\lambda_i$  depends on the  $\sigma_j$ .

The final step in the procedure is to truncate the expansion (5.13), thereby effectively relegating to  $A$  the subspace spanned by the eigenfunctions  $\psi_i$  that correspond to small eigenvalues  $\lambda_i$ . It can easily be shown that these functions contribute little to the integrals in the constraints (5.9) compared with their contribution to the sum (5.13). It is necessary to decide where (and how) the expansion is to be truncated, and here several options are available (e.g. Wiggins, 1972; Jackson, 1973). In the following section I illustrate a pragmatic approach, which appears to be reliable when systematic errors in the data can be ignored. Moreover, it is particularly useful when the magnitude of the random errors in the data are poorly estimated. It rests on the fact that as  $\lambda_i$  decreases the functions  $\psi_i$  tend to develop more and more small-scale structure and the coefficients  $\alpha_i$  increase (as  $\lambda_i^{-1/2}$ ) in a random way once they are dominated by errors. If one increases from unity the number of terms retained in the expansion (5.13), the result should first approach the correct solution. Then, once errors dominate, successive approximations diverge, exhibiting structure on smaller and smaller scales with larger and larger amplitude. The best estimate of  $\Omega$  one can obtain by the method is the function to which the expansion appears to be converging before the divergence takes over.

### Extracting localized averages

The idea of finding localized averages was introduced by Backus and Gilbert (1968), partly as a means of assessing what they call the resolving power of the data: namely, the degree to which information about  $\Omega$  that is contained in the data can be localized in space.

To begin, the constraints (5.1) are rescaled by dividing each by  $\int_0^R K_i(r)dr$ , yielding

$$w_i^* = \int_0^R K_i^*(r)\Omega(r)dr \quad , \quad (5.15)$$

where  $K_i^*$  is unimodular: i.e.

$$\int_0^R K_i^*(r)dr = 1 \quad . \quad (5.16)$$

Thus  $w_i^*$ , which is an observable quantity, is an average of  $\Omega$ . The idea now is to seek, for a certain value  $r_0$  of  $r$ , a set of coefficients  $\beta_i(r_0)$  such that in the linear combination

$$\sum_i \beta_i(r_0)w_i^* = \int_0^R D(r_0,r)\Omega(r)dr \quad (5.17)$$

the averaging kernel

$$D(r_0,r) = \sum_i \beta_i(r_0)K_i^*(r) \quad (5.18)$$

resembles a Dirac delta function centred at  $r = r_0$ . One can assess the degree of success by inspecting  $D$  to see how localized it is. A quantitative measure of the localization can be obtained from the spread, defined by Backus and Gilbert (1970) as

$$s(r_0;D) = 12 \int_0^R (r-r_0)^2 D^2(r_0,r)dr \quad , \quad (5.19)$$

which, as Backus and Gilbert show, can be large either if  $D$  is not well localized or if  $r_0$  is far from center  $\bar{r}$  of a well-localized  $D$ , where

$$\bar{r} = \frac{\int_0^R rD^2(r_0,r)dr}{\int_0^R D^2(r_0,r)dr} \quad . \quad (5.20)$$

The factor 12 is chosen in (5.19) because if  $D = \delta^{-1}$  for  $|r-r_0| < \delta/2$  and  $D = 0$  otherwise, then  $s(r_0;D) = \delta$ . For any given  $D$ ,  $s(r_0;D)$  is minimized when  $r_0 = \bar{r}$ . That minimum,  $\delta \equiv s[\bar{r};D(r_0,r)]$ , is the width of  $D$ , and measures the degree to which  $D$  resembles a delta function. Alternative measures have also been used.

The average (5.17) of  $\Omega$ , which I denote by  $\bar{\Omega}(\bar{r})$ , is a useful diagnostic of the function  $\Omega$ . Indeed, as Backus and Gilbert (1968) point out, it is only this that is in any sense determined by the data. Thus one can regard  $\bar{\Omega}$  as an estimate of  $\Omega$ , bearing in mind that it is really a smoothed version obtained by averaging over a characteristic distance  $\delta$ .

The determination of the coefficients  $\beta_i(r_0)$  is discussed by Backus and Gilbert (1968, 1970). One chooses a function  $J(r_0, r)$  which vanishes when  $r = r_0$  and increases monotonically away from  $r_0$ , and one minimizes the functional

$$\Delta \equiv \int_0^R JD^2 dr \quad , \quad (5.21)$$

subject at least to the constraint

$$\int_0^R D(r_0, r) dr = 1 \quad . \quad (5.22)$$

When  $\Delta$  is small the constraint (5.22) forces  $D$  to be small where  $J$  is large, and permits  $D$  to be large near  $r = r_0$  where  $J$  vanishes.

In an application of this procedure to artificial high-degree solar p-mode data (Gough, 1978b) it was found that the final result was not particularly sensitive to the form chosen for  $J$ , provided it rose steeply enough far from  $r_0$ . I shall not discuss that issue further here. All the examples illustrated in the following sections were computed with  $J(r_0, r) = 12(r-r_0)^2$ , which is the case that Backus and Gilbert (1970) discuss in detail when they consider erroneous data. Then

$$\Delta = s = \sum_{ij} S_{ij}^* \beta_i \beta_j \quad (5.23)$$

where

$$S_{ij}^* = 12 \int_0^R (r-r_0)^2 K_i^*(r) K_j^*(r) dr \quad . \quad (5.24)$$

In practice, errors in the data cause the minimization of  $\Delta$  subject to the constraint (5.22) not to provide a good measure  $\bar{\Omega}$  of the average of  $\Omega$ . The reason is that to obtain the most concentrated kernel  $D$  requires coefficients  $\beta_i$  with large magnitudes. The constraint (5.22) requires that  $\sum_i \beta_i = 1$ . Severe cancellations are therefore required in that sum, and in the sum on the left-hand side of equation (5.17). Cancellation in the latter does not actually take place when the data  $w_i^*$  contain errors.

Backus and Gilbert (1970) assume that an estimate  $E_{ij}$  of the covariance matrix of the errors in the data  $w_i^*$  is known, so that one can estimate the error  $\epsilon$  in  $\bar{\Omega}$ :

$$\epsilon^2 = \sum_{ij} E_{ij} \beta_i \beta_j \quad . \quad (5.25)$$

The idea then is to set a limit  $\epsilon_0$  on the error  $\epsilon$ , and minimize  $s$  subject to the constraint  $\epsilon \leq \epsilon_0$  and equation (5.22). The result depends on the choice of  $\epsilon_0$ . Alternatively one can set a limit  $s_0$  on  $s$ , and minimize  $\epsilon$  subject to  $s \leq s_0$  and the constraint (5.22).

As Backus and Gilbert prove, the two formulations are essentially equivalent. They point out that  $S_{ij}^*$  and  $E_{ij}$  are both positive definite symmetric matrices, so that  $s = s_0$  and  $\epsilon = \epsilon_0$  each define hyperellipsoids in the parameter space spanned by  $\beta_i$ . The intersections of these hyperellipsoids with the hyperplane  $\sum \beta_i = 1$  are also hyperellipsoids, which I denote by  $s^\dagger = \text{constant}$  and  $\epsilon^\dagger = \text{constant}$ . In general,  $E_{ij}$  is not a scalar multiple of  $S_{ij}^*$ , so the centres of the ellipsoids are not coincident. Consequently, for any given value  $\epsilon_1 \leq \epsilon_0$  of  $\epsilon$ ,  $s$  is minimized when the ellipsoid  $s^\dagger = \text{constant}$  ( $= s_2$ , say) is tangent to the ellipsoid  $\epsilon^\dagger = \epsilon_1$ , and the solution  $\beta_i$  is the point of contact. The smallest value,  $s_0$ , of  $s$  is obtained when that point is as close to the centre of the ellipsoid  $s^\dagger = s_2$  as is permissible, which implies that  $\epsilon_1$  is as large as is permissible: namely,  $\epsilon_1 = \epsilon_0$ . The argument is clearly symmetric;  $\epsilon$  is minimized for  $s = s_1 \leq s_0$  when the ellipsoid  $\epsilon^\dagger = \epsilon_2$  is tangent to the ellipsoid  $s^\dagger = s_1$ , and the smallest permissible value  $\epsilon_0$  of  $\epsilon$  is achieved when  $s_1 = s_0$ . The conditions for tangency are:

$$\sum_j (A_{ij}^* + \mu E_{ij}) \beta_j = v \quad , \quad (5.26)$$

$$\sum_{ij} S_{ij}^* \beta_i \beta_j = s_0 \quad , \quad (5.27)$$

$$\sum_i \beta_i = 1 \quad , \quad (5.28)$$

where  $A_{ij}^*$  is defined as in equation (5.5) with  $K_k$  replaced by  $K_k^*$ . Equations (5.26)-(5.28) are to be solved simultaneously for the coefficients  $\beta_i$  and the unknown parameters  $\mu$  and  $v$ . The solution is determined uniquely in terms of  $s_0$  with  $\mu > 0$ , provided  $s_0$  is in its allowable range. Evidently  $s_0$  cannot be less than its minimum value obtained by ignoring errors; it is also bounded above, as can be appreciated from the definition (5.19) when it is recognized that the magnitude of  $D$  is bounded for finite values of  $\epsilon$ .

In practice it is simpler to solve the problem implicitly, by choosing  $\mu$  and solving for  $\beta_i$ ,  $v$  and  $s_0$ . Backus and Gilbert set  $\mu = w \tan \theta$  and  $v = w \sec \theta$  with  $0 < \theta < \pi/2$ , and  $w > 0$  chosen, for convenience, to make  $S_{ij}^*$  and  $wE_{ij}$  of comparable numerical size. Then they introduce

$$W_{ij}(\theta) = S_{ij}^* \cos \theta + wE_{ij} \sin \theta \quad . \quad (5.29)$$

Equation (5.26) now takes the form

$$\sum_j W_{ij} \beta_j = b \quad . \quad (5.30)$$

The matrix  $W_{ij}$  is symmetric and positive definite, and therefore has a positive definite inverse  $W_{ij}^{-1}$ . Therefore for any chosen  $\theta$  equations (5.30) may be solved, whence  $s_0$  and  $\epsilon^2$  may be evaluated from equations (5.27) and (5.25).

The parameter  $\theta$  determines the extent to which errors are to be restricted, at the expense of permitting  $D$  to be less well confined. When  $\theta = 0$  the errors are ignored, and when  $\theta = \pi/2$  no attempt is made to localize  $D$ . The aim is to find a tradeoff somewhere between.

The choice of  $\theta$  is discussed by Backus and Gilbert (1970). As with the spectral method, the pragmatic approach is useful, especially when the estimate  $E_{ij}$  of the covariance matrix is uncertain. This is illustrated in the accompanying paper (Christensen-Dalsgaard and Gough, 1984a).

Examples of the application of the localized averaging procedure to real solar data are presented by Gough (1982) and Hill et al. (1984).

## 6. PHASE THREE: SOME NUMERICAL EXAMPLES

The principal object of this section is to illustrate an application of the spectral method to artificial solar data. A similar application of the localized averaging procedure is discussed in the accompanying paper (Christensen-Dalsgaard and Gough, 1984a).

The example I present in some detail concerns the determination of the horizontal component of the large-scale subphotospheric flow using high-degree  $f$  and  $p$  modes. Thus the data mimic those discussed by Deubner et al. (1979), Deubner (1983), Rhodes et al. (1983a) and Hill et al. (1983).

A model of the solar envelope was constructed in the usual way by integrating the equations of stellar structure inwards from the photosphere. The procedure was identical to that used by Hill et al. (1984). Upon this was imposed an artificial horizontal equatorial velocity  $\Omega$  in excess of the surface rotation, which is plotted in Figures 2-4. The form of the function  $\Omega$  was chosen to have several different features to test what the method could resolve: a gentle growing oscillation at low values of  $x = \log p$ , followed by three straight portions with corners, chosen such that  $\Omega$  increases at great depth. Straight portions were chosen because the  $K_i$  are not straight, and corners test the spatial resolution of the procedure. Otherwise the choice was quite arbitrary.

For the purpose of the inversion 45 five-minute sectoral modes were chosen, distributed approximately uniformly along the 9 lowest ridges in the  $k$ - $\omega$  diagram. The modes are indicated by crosses in Figure 1. The adiabatic eigenfrequencies  $\omega_0$  were computed in the usual way, and the

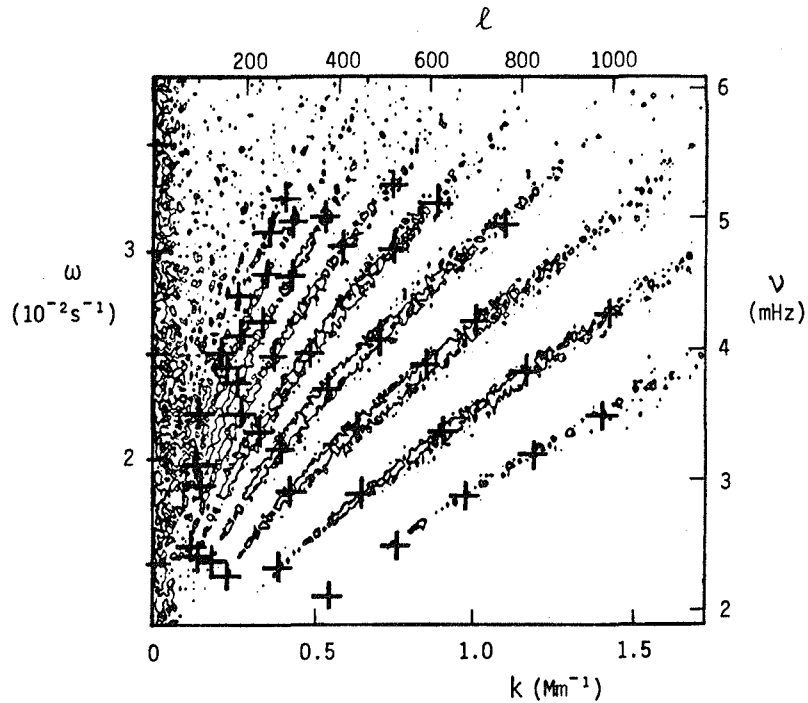


Fig. 1.  $k$ - $\omega$  diagram of prograde sectoral modes computed from the observations of Hill et al. (1983). The crosses indicate the modes that were used in the inversions displayed in Figures 2-4.

perturbations  $\omega_1$  produced by the artificial function  $\Omega$  were computed from the eigenfunctions using equation (4.5). The modes considered are all concentrated in the outer layers of the Sun; therefore it was adequate to compute only in the region  $r > 0.5 R$ .

Three inversions were performed. The first used the raw frequencies; the other two used data constructed by adding independent Gaussian-distributed noise to the frequency perturbations  $\omega_1$ . The same set of modes was used in all three cases.

In Figure 2 are displayed the expansions (5.13) truncated at four different points. They include  $I = 9, 15, 20$  and  $33$  basis functions, with eigenvalues satisfying  $\lambda_1/\lambda_1 > 3 \times 10^{-2}, 10^{-2}, 10^{-3}$  and  $10^{-8}$ . As expected, convergence generally improves as the number of modes is increased. If even more modes are included the accuracy of the expansion deteriorates slightly, presumably as a result of rounding errors. At depths greater than  $x \approx 12.5$ , the most deeply penetrating oscillation eigenfunctions and their associated kernels  $K_1$  drop to zero, and the modes cannot sense  $\Omega$ . The inferred velocity falls to zero too.

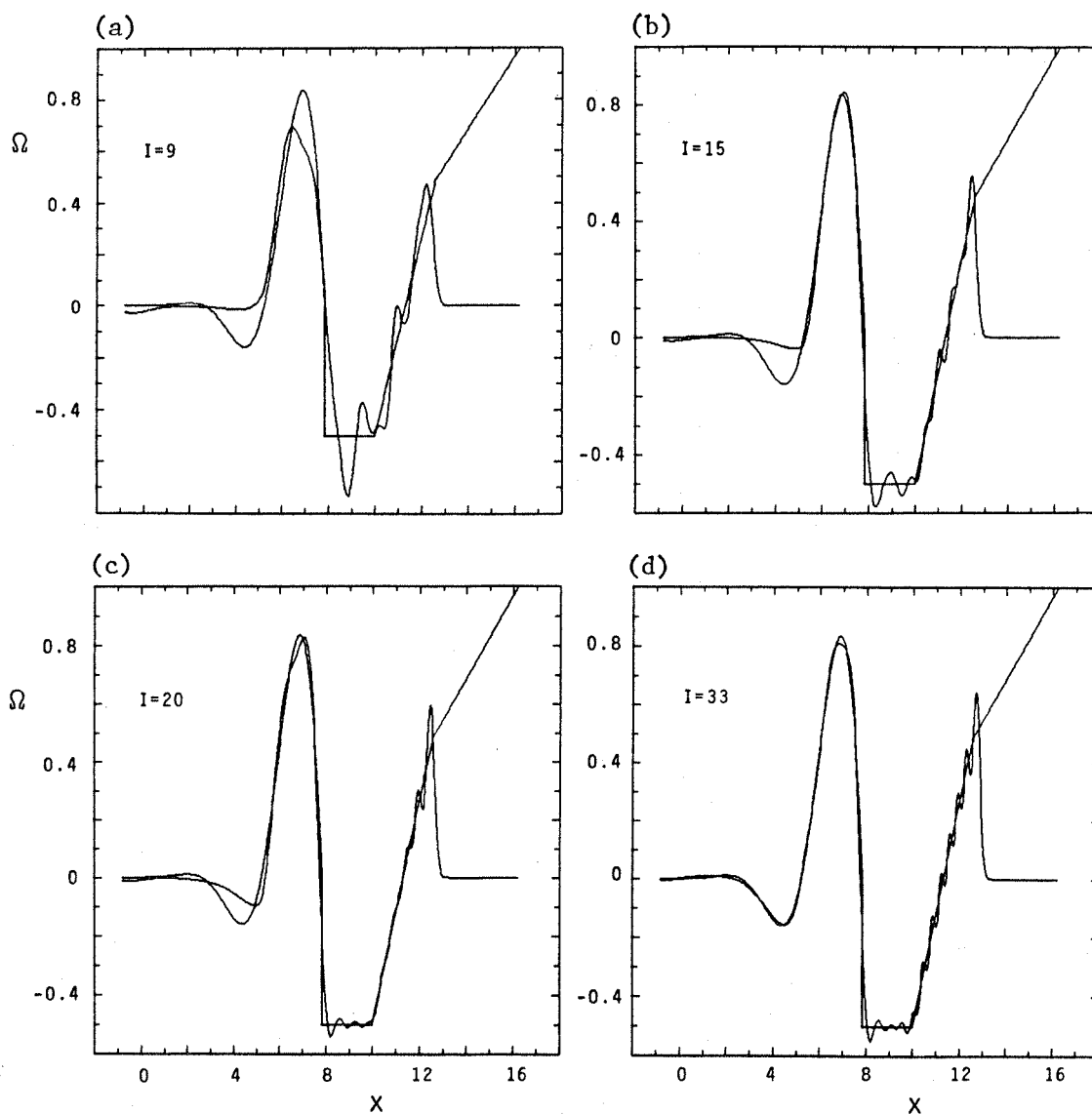


Fig. 2. Inversions of error-free frequency splittings  $\omega_{1i}$  of the 45 modes indicated in Figure 1. Both the results of the inversions and the imposed horizontal velocity  $\Omega$  from which  $\omega_{1i}$  were computed are shown; the latter is easily identifiable because it is common to all the panels. The eigenvalues  $\lambda_i$  retained in the expansions represented in panels (a)-(d) satisfy  $\lambda_i > \lambda_m$ , where  $\lambda_m/\lambda_1 = 3 \times 10^{-2}$ ,  $10^{-2}$ ,  $10^{-3}$  and  $10^{-8}$  respectively; corresponding numbers  $I$  of basis functions  $\psi_i$  are 9, 15, 20 and 33. Since the problem is linear, the ordinate scale is arbitrary. The independent variable is  $x = \log_{10} p$ .

In Figure 3 are inversions of data with errors having standard deviations of 5 per cent of the raw values. The first two inversions are hardly distinguishable from the corresponding inversions with error-free data. The errors have been averaged to some degree by the inversion procedure, as a result of combining redundant information, leading to an inferred velocity with an error generally rather less than 5 per cent of the mean magnitude of  $\Omega$ . Evidently the error that is present is a result of imperfect resolution resulting from retaining no more than 15 basis functions. When 20 basis functions are included (Figure 3c), however, the errors are beginning to become important, and erroneous small-scale structure is introduced into the inferred velocity. Presumably this would not have happened at this level of truncation had substantially more than 45 data been included in the inversion. Figure 3d includes 26 modes, which have  $\lambda_i/\lambda_1 > 10^{-5}$ ; now erroneous rapid oscillations dominate the expansion, though the general behaviour of  $\Omega$  is still discernable.

The final example, which is illustrated in Figure 4, was computed from data with 50 per cent errors. Even though the error in the inferred velocity is as great as 100 per cent in places, the inverted data still retain the basic features of  $\Omega(r)$ . When 20 basis functions are used the characteristic wiggles similar to those in Figure 3d are introduced.

The conclusion to be drawn is that the spectral method is likely to be useful for solar inversions to infer the large-scale subphotospheric velocity (including rotation).

As a secondary example I present the results of a single iteration of a density inversion. This work was carried out some years ago in collaboration with A. J. Cooper (unpublished), to ascertain whether it was likely to be possible to infer the distribution of the density of the Sun from whole-disk data taken from a spacecraft. Two simple models of the Sun were constructed, each consisting of an inner polytrope of index 3 to represent the radiative interior and a polytropic exterior of index 1.5 to represent the convection zone. The transition in one case was at  $x \equiv r/R = 0.75$ , and in the other it was at  $x = 0.80$ . Eigenfrequencies of the two models were computed for low-degree modes of relatively low order, assuming  $\gamma = 5/3$  throughout. One set was considered to be the observations; the other set was the theoretical eigenfrequencies of a trial solar model. No errors were added. Inversions of equation (4.4) were carried out using both the spectral expansion and the localized averaging procedure to try to deduce the density distribution of the model that provided the artificial data. The fact that it had been constructed from polytropes is not important since that information was not available for the inversion: only the mass, the radius and the eigenfrequencies were used. The constraint of preserving the mass is simply of the form (4.4), with the left-hand side replaced by zero and with  $S = 1$ ; it was handled in the same way as the frequency constraints.

The results of the two pairs of inversions are shown in Figures 5 and 6, where the inferred density differences  $\delta\rho$  between the trial and



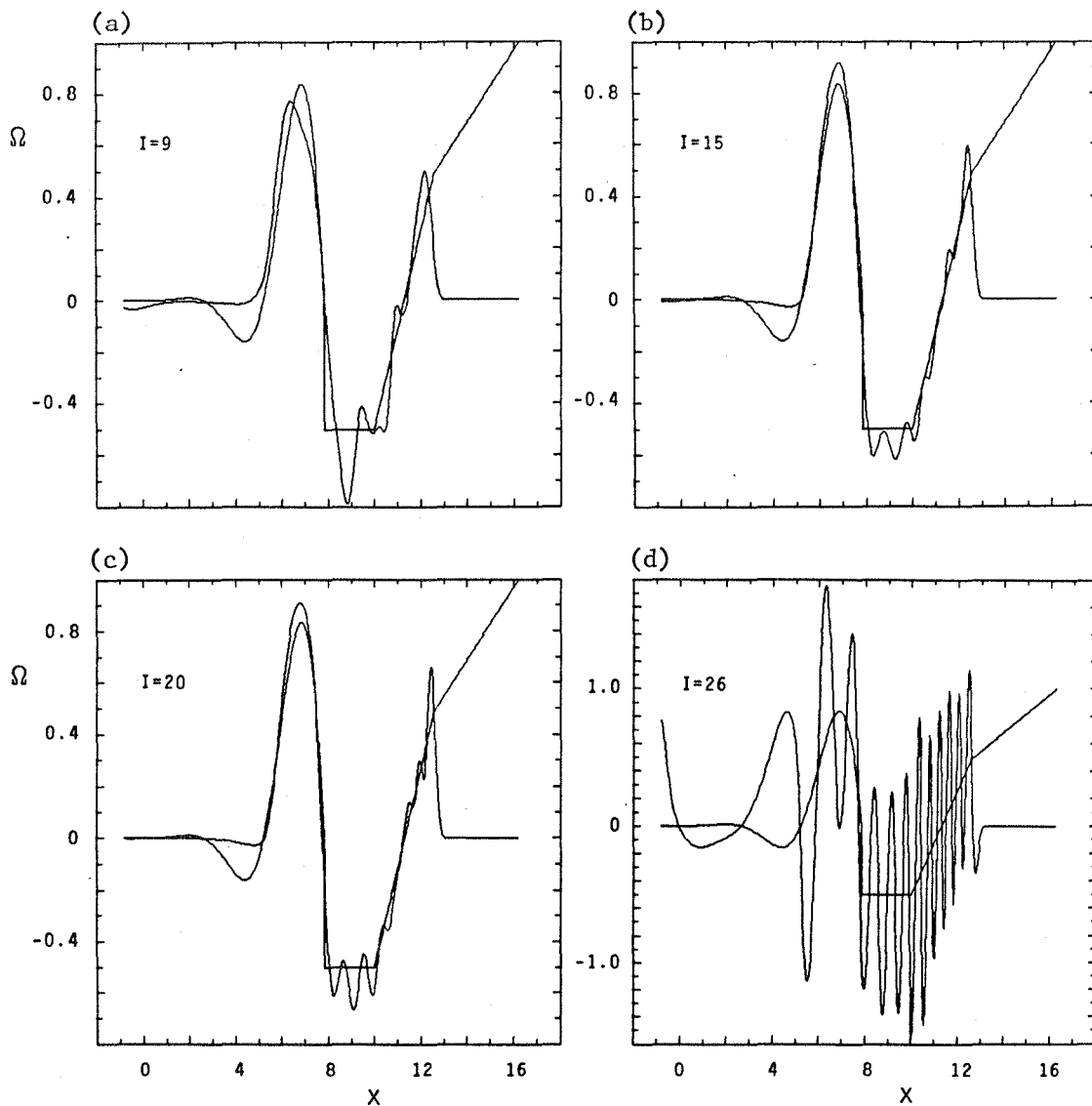


Fig. 3. Inversions of frequency splittings to which independent Gaussian distributed errors with 5% standard deviation have been added. The lower bounds  $\lambda_m$  to the eigenvalues retained in panels (a)-(d) are given by  $\lambda_m/\lambda_1 = 3 \times 10^{-2}$ ,  $10^{-2}$ ,  $10^{-3}$  and  $10^{-5}$  respectively. Note the difference in the ordinate scale in panel (d).

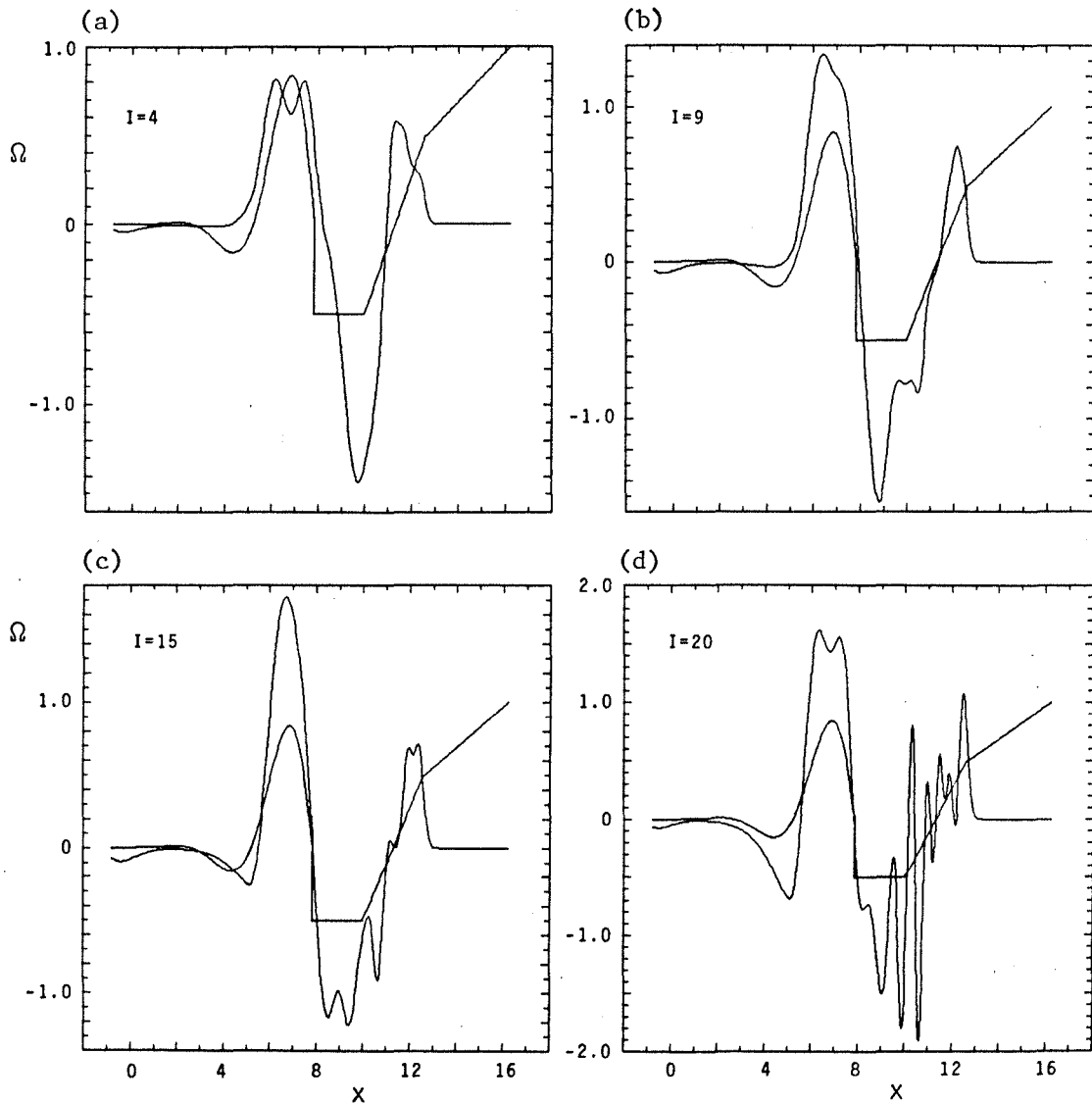


Fig. 4. Same as Figure 3, except that 50% errors have been added to the frequency splittings and that  $\lambda_m/\lambda_1 = 10^{-1}$ ,  $3 \times 10^{-2}$ ,  $10^{-2}$  and  $10^{-3}$ . The expansion with  $\lambda_m/\lambda_1 = 10^{-1}$ , which includes only four terms, is similar in form to similar truncations with 10% or no errors, but its amplitude is roughly twice as great.

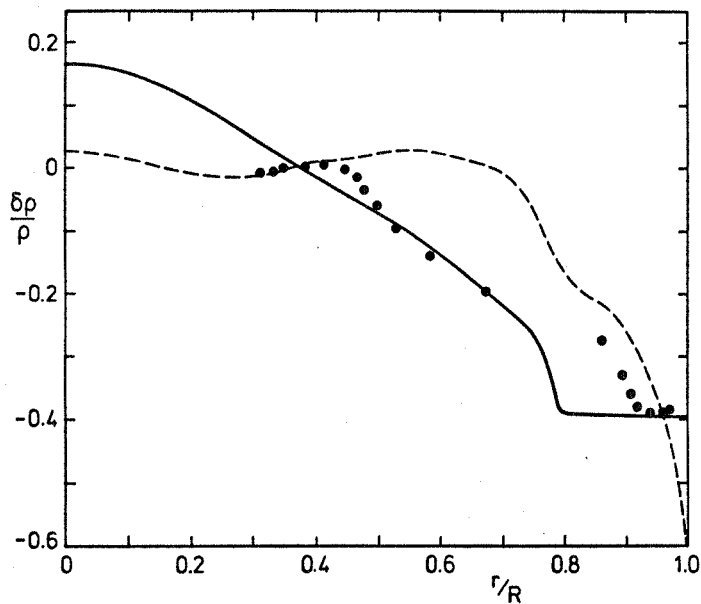


Fig. 5. The continuous line is the relative density difference  $\delta\rho/\rho$  between the 'true' artificial model of the Sun and a trial. The dashed line represents an estimate of  $\delta\rho/\rho$  from the spectral expansion obtained from a single iteration from the trial by the procedure outlined in the text. The filled circles are corresponding localized averages: no well-localized kernels could be constructed with centres at radii smaller than 0.3 R. Dipole ( $\ell = 1$ ) and quadrupole ( $\ell = 2$ ) p modes of orders 1-9 were used in both inversions.

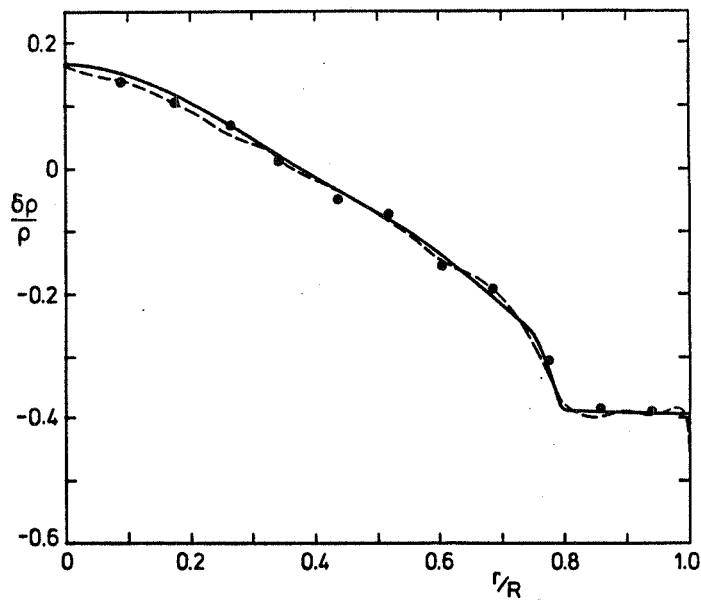


Fig. 6. Same as Figure 5, but using modes  $g_4-p_5$  ( $\ell = 1$ ) and  $g_3-p_5$  ( $\ell = 2$ ).

the artificial Sun are compared with the actual difference. Both used just 18 modes. The two inversions in Figure 5 were performed with the 9 lowest-order p modes with  $\ell = 1$  and  $\ell = 2$ . The results are not very good, though the general trend of the exact curve is exhibited by the inversions. The poor quality of the inversion would have been evident without prior knowledge of the correct  $\delta\rho$ , partly because in most places it was not possible to produce nicely localized optimal kernels D, particularly near the centre of the star, and partly because on the whole the two inversion techniques gave different results.

When some of the p modes are replaced by the quadrupole f mode and some low-order g modes, the situation is quite different. Now sharply peaked kernels D can be produced, and the two methods both yield very similar and quite accurate results. These are shown in Figure 6.

We have also performed some experiments with other groups of modes, and with data containing errors. The results of that investigation will be reported elsewhere.

## 7. OTHER INVERSION METHODS

Other criteria can be considered for selecting from the infinity of functions  $\Omega$  that satisfy the constraints (5.1). For example, one can choose the smoothest or the flattest function by minimizing

$$\int_0^R \left( \frac{d^2\Omega}{dr^2} \right)^2 dr \quad \text{or} \quad \int_0^R \left( \frac{d\Omega}{dr} \right)^2 dr \quad (7.1)$$

subject to the constraints (5.1). The latter has been used by Gough (1982, 1984) with the solar data reported by Hill et al. (1982) and Scherrer and Delache (1984). In view of the fact that functions in the annihilator A of (5.1) tend to be rapidly varying, one would expect these procedures to give results similar to the spectral expansion.

Another procedure that I must mention is that described by Ulrich et al. (1979). It was the first to be used on real solar data (Deubner et al. 1979). Despite the fact that it was demonstrated to be internally inconsistent, yielding unfaithful results in a particular artificial case (Gough, 1978b), it behooves me to discuss it further because Rhodes et al. (1983b) plan to use it again on future observations to measure the solar rotation.

The procedure appears to be based on the incorrect assumption that the kernels  $K_i^*$  are highly localized. The essence was to replace equation (5.1) by

$$w_i = \Omega(r_{1i}) \quad , \quad (7.2)$$

where  $r_{1i}$  measured the "effective depth" of  $K_i^*$ . The definition adopted for  $r_{1i}$  can easily be shown to be equivalent to the first moment of  $K_i^*$ , though Ulrich et al. (1979) did not describe it in those terms. As was

pointed out by Gough (1978b), the procedure would be adequate if  $\Omega$  were a linear function of  $r$ , for then equation (7.2) is correct even though  $K_i^*$  is not localized. Nevertheless, Deubner et al. found empirically that the  $\Omega$  is not a linear function. Subsequent observations by Rhodes et al. (1983a) did not support the finding that  $\Omega$  varies with depth, which led Rhodes et al. (1983b) to suggest that inversion procedures of the type described here are not necessary. However, the more extensive observations of Hill et al. (1984) suggest that  $\Omega$  does vary, both in space and in time.

In support of the continued use of equation (7.2) Rhodes et al. (1983b) argue that it can be regarded as containing just the first two terms of a Taylor expansion of  $\Omega$ . If that were so, Deubner et al. (1979) would have represented  $\Omega$  by a linear function of depth. Rhodes et al. (1983b) went on to say that, should future observations merit, higher terms could be included, after which a least-squares analysis could be carried out. Thus one would set

$$\Omega \approx \sum_k \gamma_k z^k \quad (7.3)$$

where  $z = R-r$ , and determine  $\gamma_k$  by minimizing

$$\chi^2 = \sum_i \mu_i \left[ \sum_k \gamma_k (z_{ik})^k - w_i \right]^2, \quad (7.4)$$

where

$$z_{ik} = \left( \int_0^R z^k K_i^* dz \right)^{1/k}, \quad (7.5)$$

which are essentially the higher-order effective depths of Ulrich et al. (1979); the coefficients  $\mu_i$  are constants that weight the data according to the errors (e.g.  $\mu_i = \sigma_i^{-2}$ ).

This procedure has already been carried out (Gough, 1982) on the rotational splitting data of Hill et al. (1982) and Claverie et al. (1981) using polynomials of up to eighth degree, and I have shown (unpublished) it to work quite well on artificial data (when it can be reliably tested) provided the assumed variation of  $\Omega$  is on a scale large enough to be resolved by a modest number of terms. However, it does not work when  $\Omega$  varies on a smaller scale. This is illustrated in Figure 7, where polynomials of up to sixth degree are fitted to the data that were inverted by the spectral method to produce Figures 2 and 3. None of the polynomial approximations fits the original curve particularly well. Moreover, without a knowledge of the original function  $\Omega$  from which the  $w_i$  were computed, there is no way of knowing where the polynomial approximations are least good. The failure to converge, however, does indicate that the results cannot be trusted. If the degree of the polynomial is increased further, or if substantially larger errors are added to the data, the mismatch is aggravated severely. Of course, one can always find a smooth function whose values are arbitrarily close to any angular velocity  $\Omega$  for which the radius of convergence of a power series

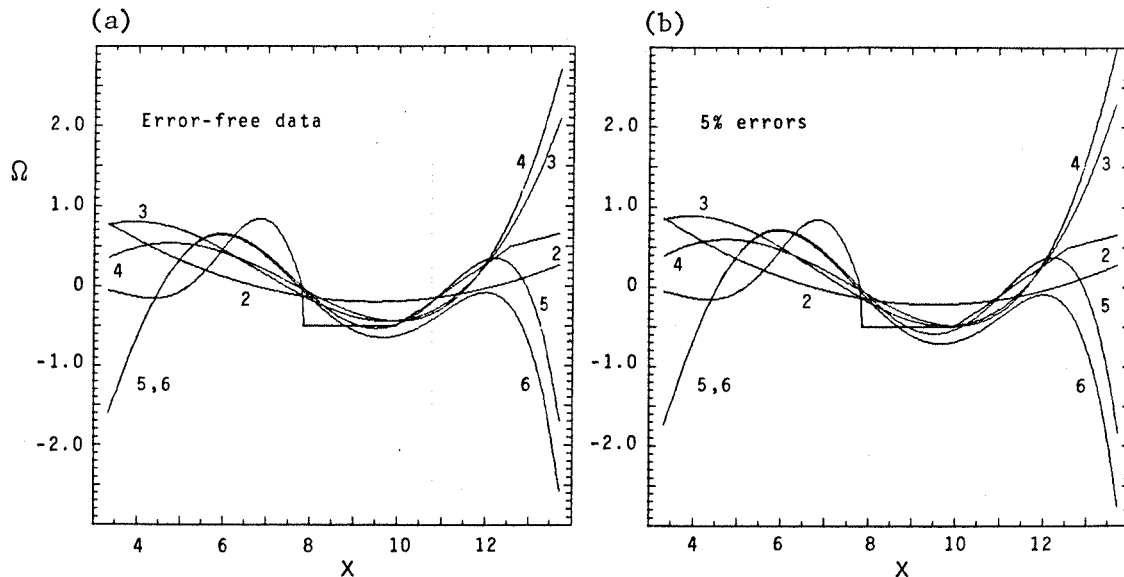


Fig. 7. Polynomial approximations to the same horizontal velocity  $\Omega$  as was used in the inversions in Figures 2-4, computed by  $\chi^2$  minimization as described in the text. Panels (a) and (b) use the same data as were used for Figures 2 and 3 respectively. The degrees of the polynomials are indicated in the figure.

representation is greater than the range of  $r$  over which  $\Omega$  can be measured, and so one might think that this method could be used with many more terms. However, the procedure is then unstable to small errors in the data, and in any case the method is impractical because the matrix that must be inverted to compute the coefficients of the power series is ill conditioned.

My conclusion is that there is little point in using power series approximations in cases when the constraints can be written as linear integral equations of the type (5.1). However, in more complicated circumstances it may not be unwise to try it, because the expansion is usually simple to carry out. If one is fortunate enough to obtain a stable apparently convergent representation, then the problem is probably solved. But if the problem is expressible in the form (5.1), it seems to be expedient to expand in the more natural basis functions  $\psi_1$ .

## 8. DISCUSSION

The conclusion to be drawn from this and the accompanying article (Christensen-Dalgaard and Gough, 1984a) is that both the Gilbert-Backus optimal localized averaging procedure and the spectral expansion are likely to be useful for inverting solar oscillation frequencies. It has been shown that the method should certainly work given accurate knowledge of the rotational splitting of modes that have already been

detected observationally, and evidence has been presented to suggest that it should work for inferring the Sun's pressure and density distribution too.

I should stress however, that the more primitive assessments of Phase Two of the inversion will continue to have some value. I have already described how model-fitting calculations have given us estimates of the Sun's initial helium abundance and the present depth of the convection zone, and how an appreciation of the asymptotic structure of high-frequency p modes has permitted us to infer that at least part of the cause of the unresolved discrepancy between theory and observation lies in the convection zone, or near its base. It is in the convection zone that there are substantial uncertainties in solving the forward problem, for there nonadiabatic effects and Reynolds stresses may play a significant role. It is also in the convection zone that the equation of state appears to be most uncertain. Shibahashi *et al.* (1983, 1984) and Ulrich (1982) have studied the influence of electrostatic interactions on the equation of state, using the Planck-Larkin cutoff, and have assessed how they effect the eigenfrequencies (see Ulrich and Rhodes, 1983). Christensen-Dalsgaard and Gough (unpublished) have made similar studies, using variants of the theory used by Fontaine *et al.* (1977); internal partition functions computed both from the static screened Coulomb potential, as was first used in oscillation calculations by Berthomieu *et al.* (1980), and from the confined atom model, using the programme described by Däppen (1980), together with estimates of the additional degree of ionization induced by time-dependent perturbations to the bound-state potential and the influence of neighbouring neutral species, have all been considered. In no case has adequate agreement between observation and theory been achieved. Whether this implies that current procedures for calculating the equation of state are woefully inadequate or whether it implies that there is something else wrong with our models of the Sun is yet undetermined. It may be that an application of the inversion procedures of Phase Three will give us the next clue.

It may also require Phase Three methods to resolve the apparent discrepancy raised by Scherrer and Delache (1984) and van der Raay *et al.* (1984), who claim respectively that the mean normalized period interval

$$\overline{(P_{n+1,\ell} - P_{n,\ell})} / [\ell(\ell+1)]^{1/2} ,$$

(where  $P_{n,\ell}$  is the period of the mode of order  $n$  and degree  $\ell$ ) between dipole and quadrupole g modes in the period range 3-5 hr is 38.6 min and 41.2 min, whereas the mean theoretical separation of the corresponding modes of Christensen-Dalsgaard's standard Model 1 of the Sun is only 34.5 min. This discrepancy may not be real, however, because the oscillations observed may have been misidentified.

Formal inversions of the limited rotational splitting data from low-degree modes that have appeared in the literature (Claverie, *et al.* 1981; Hill *et al.* 1982; Scherrer and Delache, 1984) have already been

carried out (Gough, 1982, 1984; Hill et al. 1982; Campbell et al., 1983), but, even if the data have been correctly interpreted, they are too sparse for a well resolved angular velocity to be inferred. This has led to some controversy over the value of the gravitational quadrupole moment  $J_2$  of the Sun that is induced by the centrifugal forces, and its implications regarding tests of theories of gravitation, such as general relativity, from the precession of the perihelion of Mercury. It is undeniable that the splitting data, coupled with the planetary orbit data, do not contradict general relativity (Gough, 1982; Campbell et al. 1983). However, both H. Hill et al. (1982) and Campbell et al. (1983), using a particular kind of coarsely resolved analysis, have concluded that  $J_2$  is substantially greater than the minimum value permitted by the data, thereby challenging the analysis of the orbit data or supporting alternative theories of gravitation. The estimate  $\bar{\Omega}$  of the localized averaging procedure, however, is not in conflict; moreover the resolving width is not everywhere smaller than the scale of variation of the  $J_2$  kernel  $J$  [see Figure 13 of Christensen-Dalsgaard and Gough (1984a)], and  $J_2[\bar{\Omega}]$  does not necessarily underestimate the true  $J_2$  (Gough, 1982). It is important to realize that the conclusion that there is a conflict rests crucially on an additional assumption, which is not suggested by the data and which results in both the restriction of the space  $O$  of acceptable functions  $\Omega$  by the imposition of unsubstantiated (and incompletely defined) constraints and the inclusion of components from the annihilator  $A$ . This would be a perfectly acceptable procedure if the restrictions on  $\Omega$  were stated clearly, for then the plausibility of the conclusion could more easily be judged by the casual reader. An advantage of an inversion procedure such as the spectral expansion is that the truncation of (5.13) (or the inclusion of functions from  $A$ ) makes one acutely aware of the assumptions one has made. What is really necessary, however, is a much larger data set, with substantially greater resolving power.

An inversion of frequency splitting of high-degree modes observed by F. Hill et al. (1983) has been undertaken by F. Hill et al. (1984), using the localized averaging technique of Backus and Gilbert (1970). The results suggest variations in the large-scale subphotospheric velocity that may be associated with giant convective cells. Aside from using asymptotic approximations to the frequencies (Gough, 1984), no (Phase Three) inversions of real solar data have yet been performed to obtain the density stratification of the Sun.

What are the differences between the two principal inversion methods I have discussed? Were the data to be free from error, the minimum value of  $\lambda_1$  that is retained in spectral expansions can, at least in principle, be adjusted to make the solution satisfy the constraints (5.1) as accurately as one pleases. This is not so of the localized averaging technique, for even if errors are ignored, there is always a minimum resolution width of the optimal kernels. Therefore adopting  $\bar{\Omega}(\bar{r})$  as a representation of  $\Omega(r)$  can never be exact. For this reason some geophysicists (e.g. Parker, 1977a) prefer not to use the localized averages for inversions, favouring simply a particular example, such as is provided by the spectral expansion, of the infinity of solutions that



satisfy the constraints (5.1) within the limits set by the precision of the data. As is evident from the examples illustrated in Figures 2-4, those solutions have unreal small-scale structure. In interpreting the spectral representations one's eye should smooth out the wiggles, by imagining an average over the resolving width of the optimal localized kernels. That average is provided by  $\bar{\Omega}(\bar{r})$ , however, which approximates the average of all acceptable solutions to the constraints (5.1), including those that have components in the annihilator A. It is important to realize that A is determined by the limitations of the observations, and has no other physical significance. Therefore, as Backus and Gilbert (1967) point out, it is only  $\bar{\Omega}$  that is a valid representation of the true function  $\Omega$ , though it is necessary to recognize that  $\Omega$  may have small-scale structure not present in  $\bar{\Omega}$  that cannot be resolved by the data. What the representation  $\bar{\Omega}$  means can be assessed by inspecting the optimized kernels D.

One should appreciate also that it can be dangerous to use  $\bar{\Omega}$  for estimating quantities that are nonlinear functions of  $\Omega$ . An example is the gravitational quadrupole moment  $J_2$ , which is an integral of  $\Omega^2$  weighted with the nonnegative function  $J$  displayed in Figure 13 of Christensen-Dalsgaard and Gough (1984a). The rectification of any small-scale oscillatory component of  $\Omega$  that is not present in  $\bar{\Omega}$  adds a positive contribution to the estimate  $J_2[\bar{\Omega}]$  of the moment, which tends to make  $J_2[\bar{\Omega}]$  an underestimate. In addition the smoothing produced by the averaging causes  $\bar{\Omega}$  to be systematically above or below the true functions  $\Omega$  in regions where the curvature of  $\Omega$  is positive or negative.

An issue that to some is of considerable importance concerns the computing resources required to carry out the two techniques. The computation of the spectral expansion itself is dominated by calculating the large matrix  $A_{ij}$  and, to a lesser extent, finding its eigenvectors and eigenvalues for constructing the matrix  $U_{ij}$ . For each inversion, this is done once. In contrast, the construction of the localized kernels  $D(r_0, r)$  involves a similar amount of work to that required for the entire spectral expansion at each point  $r_0$ . Strictly speaking one should compute D to determine the resolving power of the data, in order to gauge the reliability of the spectral expansion. But since that takes much longer than the expansion itself, one might be tempted not to.

Another difference between the two techniques is that the spectral expansion always provides a solution, whereas the localized averaging technique may not. The solutions by spectral expansion in Figures 2-4 extend over the entire domain, including the region  $x > 12.5$  about which the data contain no useful information and where the spectral expansion is incorrect. Of course one could have easily inferred that the expansion cannot be trusted at great depths by inspecting the kernels  $K_i^j$  and noticing that there they are all zero. However, it is less obvious that the representation is poor when there is a region where all the kernels have a similar functional form but are not zero. It is for these regions particularly that the resolving power must be established. Thus one might consider an apparent advantage of the optimal localized averaging procedure to be that it does not always work, either by

failing to provide a sharply localized kernel centred about a desired value of  $r$ , or even by failing to produce any suitable kernel at all. Of course a meticulous worker, not hampered by limitations of computer resources, would certainly compute the resolving power, in which case the advantage would not be real. However, it is then a trivial matter to compute  $\bar{\Omega}$ , so it seems unwise not to do so. It is evident that the spectral method and the averaging technique complement one another, to the extent that it would often be expedient to use both.

Finally I must mention that it is important to be able to identify in the observations which modes correspond to the oscillation frequencies that have been measured. In the past the identification has often been made by indirect methods, and indeed that must always be so to some extent because the order  $n$  can never be measured directly. We know that in some cases, however, it is not necessary to have a precise identification. A systematic study of when that is so has never been undertaken.

#### ACKNOWLEDGMENTS

The splitting kernels  $K_l$  used in the spectral expansion reported in this article were taken from the set used by Hill, Gough and Toomre (1984) to invert real solar data. The density inversions were carried out in collaboration with A. J. Cooper. I am very grateful to A. J. Cooper, F. Hill and J. Toomre, and to J. Christensen-Dalsgaard, A. Dziemowski and K. Whaler, for useful discussions.

#### REFERENCES

- Ando, H. and Osaki, Y. 1976: Publ. Astron. Soc. Japan, 27, 581-603  
Backus, G. and Gilbert, F. 1967: Geophys. J. R. astr. Soc., 13, 247-276  
Backus, G. and Gilbert, F. 1968: Geophys. J. R. astr. Soc., 16, 169-205  
Backus, G. and Gilbert, F. 1970: Phil. Trans., 266A, 123-192.  
Berthomieu, G., Cooper, A. J., Gough, D. O., Osaki, Y., Provost, J. and Rocca, A. 1980: in Nonradial and nonlinear stellar pulsations (eds. H. A. Hill and W. A. Dziemowski, Springer, Heidelberg), pp. 307-312  
Campbell, L., McDow, J. C., Moffatt, J. W. and Vincent, D. 1983: Nature, 305, 508-510  
Christensen-Dalsgaard, J., Duvall Jr, T. L., Gough, D. O. and Harvey, J. W. 1984: in press  
Christensen-Dalsgaard, J. and Gough, D. O. 1980: Nature, 288, 544-547  
Christensen-Dalsgaard, J. and Gough, D. O. 1981: Astron. Astrophys., 104, 173-176  
Christensen-Dalsgaard, J. and Gough, D. O. 1984a: 'Rotational inversion from global solar oscillations,' these proceedings  
Christensen-Dalsgaard, J. and Gough, D. O. 1984b: 'Implications of observed frequencies of solar p modes,' these proceedings

- Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. and Roca Cortez, T. 1981: Nature, 293, 443-445
- Cox, J. P. 1980: Theory of Stellar Pulsation (Princeton Univ. Press)
- Däppen, W. 1980: Astron. Astrophys., 91, 212-220
- Deubner, F-L. 1975: Astron. Astrophys., 44, 371-375
- Deubner, F-L. 1983: Solar Phys., 82, 103-109
- Deubner, F-L., Ulrich, R. K. and Rhodes Jr , E. J. 1979: Astron. Astrophys., 72, 177-185
- Duvall Jr , T. L. 1982: Nature, 300, 242-243
- Duvall Jr , T. L. and Harvey, J. W. 1983: Nature, 302, 24-27
- Dziembowski, W. A. and Goode, P. R. 1984: Mem. Soc. astr. Italiana, in press
- Fontaine, G., Graboske Jr , H. C. and Van Horn, H. M. 1977: Astrophys. J. Suppl., 35, 293-358
- Gough, D. O. 1977: in The energy balance and hydrodynamics of the solar chromosphere and corona (eds. R. M. Bonnet and P. Delache, G. de Bussac, Clairmont-Ferrand), pp. 3-36
- Gough, D. O. 1978a: in Pleins feux sur la physique solaire (ed. J. Rösch, CNRS, Paris), pp. 81-103
- Gough, D. O. 1978b: in Proc. workshop on solar rotation (eds. G. Belvedere and L. Paternò, Univ. Catania Press), pp. 255-268
- Gough, D. O. 1981: Mon. Not. R. astr. Soc., 196, 731-745
- Gough, D. O. 1982: Nature, 298, 334-339
- Gough, D. O. 1983a: in Pulsations in classical and cataclysmic variables (eds. J. P. Cox and C. J. Hansen, JILA, Boulder), pp. 117-137.
- Gough, D. O. 1983b: in Proc. primordial helium workshop (eds. P. A. Shaver, D. Kunth and K. Kjär, European Southern Observatory, Garching), pp. 117-136
- Gough, D. O. 1984: Mem. Soc. astr. Italiana, in press
- Gough, D. O. and Taylor, P. P. 1984: Mem. Soc. astr. Italiana, in press
- Gough, D. O. and Toomre, J. 1983: Solar Phys., 82, 401-410
- Hansen, C. J., Cox, J. P. and Van Horn, J. H. 1977: Astrophys. J., 217, 151-159
- Harvey, J. W. and Duvall Jr , T. L. 1984: these proceedings
- Hill, F., Gough, D. O. and Toomre, J. 1984: Mem. Soc. astr. Italiana, in press
- Hill, F., Toomre, J. and November, L. J. 1983: Solar Phys., 82, 411-425
- Hill, H. A., Bos, R. J. and Goode, P. R. 1982: Phys. Rev. Lett., 49, 1794-1797
- Jackson, D. D. 1972: Geophys. J. R. astr. Soc., 28, 97-109
- Jackson, D. D. 1973: Geophys. J. R. astr. Soc., 35, 121-136
- Lamb, H. 1932: Hydrodynamics (Cambridge Univ. Press)
- Lanczos, C. 1961: Linear differential equations (Van Nostrand, London)
- Ledoux, P. and Perdang, J. 1980: Bull. Soc. Math. Belg., 32, 133-159
- Ledoux, P. and Walraven, T. 1958: Handb. Physik, 51, 353
- Lubow, S. H., Rhodes Jr , E. J. and Ulrich, R. K. 1980: in Nonradial and nonlinear stellar pulsations (eds. H. A. Hill and W. A. Dziembowski, Springer, Heidelberg), pp. 300-306
- Lynden-Bell, D. and Ostriker, J. P. 1967: Mon. Not. R. astr. Soc., 136, 293-310

- Parker, R. L. 1977a: Ann. Rev. Earth Planet. Sci., 5, 35-64
- Parker, R. L. 1977b: Rev. Geophys. Space Phys., 15, 446-456
- van der Raay, H. B., Claverie, A., Isaak, G. R., McLeod, C. P. and Roca Cortez, T. 1984: Mem. Soc. astr. Italiana, in press
- Rhodes Jr, E. J., Harvey, J. W. and Duvall Jr, T. L. 1983a: Solar Phys., 82, 111
- Rhodes Jr, E. J., Howard, R. F., Ulrich, R. K. and Smith, E. J. 1983b: Solar Phys., 82, 245-258
- Sabatier, P. C. 1977: J. Geophys., 43, 115-137
- Sabatier, P. C. 1978: J. Math. Phys., 19, 2410-2425
- Scherrer, P. and Delache, P. 1984: Mem. Soc. astr. Italiana, in press
- Shibahashi, H., Noels, A. and Gabriel, M. 1983: Astron. Astrophys., 123, 283-288
- Shibahashi, H., Noels, A. and Gabriel, M. 1984: Mem. Soc. astr. Italiana, in press
- Stein, R. F. and Leibacher, J. W. 1974: Ann. Rev. Astron. Astrophys., 12, 407-435
- Tassoul, M. 1980: Astrophys. J. Suppl., 43, 469-490
- Ulrich, R. K. 1970: Astrophys. J., 162, 993-1002
- Ulrich, R. K. 1982: Astrophys. J., 258, 404-413
- Ulrich, R. K. and Rhodes Jr, E. J. 1977: Astrophys. J., 218, 521-529
- Ulrich, R. K. and Rhodes Jr, E. J. 1983: Astrophys. J., 265, 551-563
- Ulrich, R. K., Rhodes Jr., E. J. and Deubner, F-L. 1979: Astrophys. J., 227, 638-644
- Unno, W., Osaki, Y., Ando, H. and Shibahashi, H. 1979: Nonradial oscillations of stars (Univ. Tokyo Press)
- Vandakurov, Yu.V. 1967: Astron. Zh., 44, 786-797
- Wiggins, R. A. 1972: Rev. Geophys. Space Phys., 10, 251-285
- Zahn, J-P. 1970: Astron. Astrophys., 4, 452-461

## ROTATIONAL INVERSION FROM GLOBAL SOLAR OSCILLATIONS

Jørgen Christensen-Dalsgaard  
High Altitude Observatory, Boulder, Colorado, and  
NORDITA, Blegdamsvej 17, København Ø

Douglas Gough  
Institute of Astronomy and Department of Applied Mathematics  
and Theoretical Physics, University of Cambridge, and Joint  
Institute for Laboratory Astrophysics, University of Colorado  
and National Bureau of Standards, Boulder, Colorado

We investigate the degree to which various sets of solar oscillations can resolve the solar internal rotation. Genuine observations were simulated by the following procedure: First an artificial angular velocity was invented by one of us, and from it the rotational splitting of a set of normal modes was calculated; to that was added some random noise. The result was treated as artificial data by the other author, acting as an observer, who attempted to recover the rotation law by using the Backus-Gilbert optimal averaging procedure. The observer knew neither the original rotation law nor the amount of noise that had been added. Finally his conclusion was compared with the actual artificial angular velocity.

### INTRODUCTION

The purpose of this investigation was to estimate how well the spherically symmetrical component  $\Omega(r)$  of the Sun's angular velocity might be inferred from rotational splitting data.

The three stages of the investigation are quite separate; the first was carried out by one of us (Author A) and the second by the other (Author B). At the time when Author A presented the first stage at the Workshop he was unaware of the conclusions of Author B. Author B was still unaware of both the artificial rotation law and the level of noise that had been added to the data at the time he presented the second stage. Author A presented the third stage by laying the actual rotation curve over the curve that had been deduced by Author B; the two authors had previously agreed on a common scale (determined by Author B) so that a direct comparison was possible. The overlaid diagram is presented here in the section on the third stage as Figure 14. Author A also announced the standard deviation  $\sigma$  of the errors he had introduced, and the gravitational quadrupole moment  $J_2$  of the solar model.

FIRST STAGE: CREATION OF THE DATA

With respect to an inertial frame of reference, solar Model 1 of Christensen-Dalsgaard (1982) was endowed with an artificial angular velocity  $\Omega(r)$ , which was taken to be independent of latitude. Linearized adiabatic oscillations of the nonrotating model were computed, and from each eigenfunction the sidereal frequency splitting  $\omega_1$  was computed as a perturbation using equation (4.5) of the accompanying paper (Gough, 1984). To these were added independent Gaussian distributed errors, with variance  $\sigma^2$ .

The modes were then divided into several sets, five of which are summarized in Table 1. The first set, designated ABS, contains the five-minute nonradial modes with degree  $\ell \leq 5$ . They represent the information that one might anticipate obtaining from whole-disk measurements such as were obtained at Antarctica (Grec *et al.*, 1980, 1983) or by the Birmingham group (Claverie *et al.*, 1980, 1981, 1982) and from Stanford (Scherrer *et al.*, 1982, 1983). The second set, DH, is a subset of the modes detected by Duvall and Harvey (1983) and Harvey and Duvall (1984). Set WDS contains modes with  $\ell \leq 3$ , as one might expect to measure from whole-disk observations from space: it includes all the p modes with frequencies up to 3.8 mHz. In addition, it includes the f modes and a few low-order g modes. The lower frequency limit was estimated from the level of sensitivity quoted in the European Space Agency's recent report on the Phase A study of DISCO, coupled with amplitude estimates obtained from the simplified version of the formalism of Goldreich and Keeley (1977) that was used by Gough (1980) and Christensen-Dalsgaard and Frandsen (1983). Set A, which was used for most of the experimentation described in the following section, contains the set WDS, the five-minute modes with  $\ell \leq 10$ , the dipole and quadrupole g

TABLE 1. Modes included in the inversions

Designation	Frequency Range* ( $\mu\text{Hz}$ )	$\ell$
ABS	2300-3800	1-5
DH	2300-3800	1-6, 8, 10, 14, 20, 24, 30, 40, 50, 60, 70, 80, 90, 100
WDS	360-3800	1-3
A	50-100 260-3800 2300-3800	1, 2 1-3 4-10
	H	
B	50-100 DH + WDS + H	1, 2

\*The symbol H refers to the seven modes reported by Hill *et al.* (1982).

modes in a frequency range comparable with that investigated by Scherrer and Delache (1984), and the seven modes (constituting subset H) reported by Hill et al. (1982). The last and largest set, B, contains all the modes in DH, WDS and H, together with the g modes from set A.

It should be emphasized that although many of the modes discussed in this paper have been detected observationally, rotational splitting has yet been reported in only a very few. Moreover, the rotational splitting computed from the artificial rotation law is not the same as has been reported from observation.

## SECOND STAGE: INVERSION OF THE DATA

Inversions were carried out using the optimal localized averaging procedure of Backus and Gilbert (1970) in the form described in the accompanying paper (Gough, 1984). To carry out the inversions it was necessary to specify the covariance matrix of the errors, which was not possible a priori. However, inspection by eye of the splitting data  $\omega_{li}$  clearly showed systematic trends, so it was inferred that the errors were small. Therefore,  $E_{ij}$  was set somewhat arbitrarily to  $10^{-2} \delta_{ij}$ , where  $\delta_{ij}$  is the Kronecker delta. An improved estimate was obtained later. The factor  $w$  in equation (5.29) of Gough (1984) was set to unity, making no attempt to balance the terms  $S_{ij}^*$  and  $wE_{ij}$ . Consequently, the inversion that was judged to be optimal did not correspond to a value of  $\theta$  that is of order unity.

In addition to applying the splitting constraints provided by the data, the inversions were constrained to satisfy the observed sidereal equatorial surface angular velocity,  $\Omega_s = 2.86 \times 10^{-6} \text{ s}^{-1}$ . This was achieved by adding to the splitting constraints (5.1) of Gough (1984) a similar integral constraint with a delta-function kernel at  $r = R$  and with  $w_i = \Omega_s$ .

Initially, inversions were carried out with 15 uniformly distributed values of  $r_0$  about which it was tried to centre the optimal averaging kernels  $D$  (see Gough, 1984). As the structure of  $\Omega$  emerged, the number of points  $r_0$  to be sampled was increased to 32, and they were distributed nonuniformly, being most closely spaced where  $\Omega$  appeared to be varying the most rapidly. Before discussing the inversions it is useful to consider the splitting kernels  $K_i^*$ . Then it is easier to appreciate the results.

In Figure 1 are displayed three typical five-minute kernels, having very similar frequencies  $\omega_0$ . Near the surface the kernels appear to be identical; there the modes do not sense the value of  $\ell$ . They penetrate to different depths, however, which are determined by the condition that the sound speed  $c$  is roughly equal to  $\omega_0 r/L$ , where  $L^2 = \ell(\ell+1)$  and  $r$  is distance from the centre of the Sun. The most striking feature of the kernels is that their amplitudes are greatest near the surface.

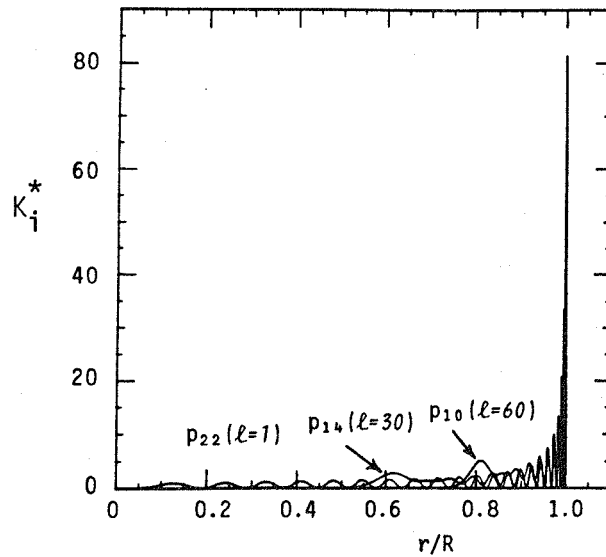


Fig. 1. Rotational splitting kernels  $K_i^*$  for three five-minute modes. The modes are  $p_{22}(\ell=1)$ ,  $p_{14}(\ell=30)$  and  $p_{10}(\ell=60)$ , having cyclic frequencies 3.24, 3.25 and 3.23 mHz, respectively.

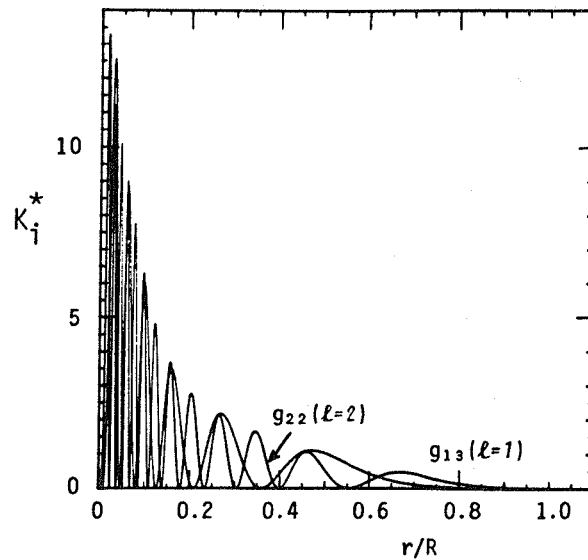


Fig. 2. Rotational splitting kernels for  $g_{13}(\ell=1)$  and  $g_{22}(\ell=2)$ , which have cyclic frequencies 48.3 and 49.4  $\mu$ Hz.



Figure 2 illustrates kernels associated with two high-order g modes with frequencies near 50  $\mu\text{Hz}$ . The modes are essentially confined beneath the convection zone, and have their greatest amplitudes at low values of  $r$ . Therefore, one would anticipate that they would be very useful for measuring  $\Omega$  close to the centre of the Sun.

Kernels of lower-order modes are smoother; a few are illustrated by Gough (1982) and Hill et al. (1982). They retain the tendency to be concentrated near the center if they are g modes and near the surface if they are p modes. On the whole the f modes of high degree are confined to the surface layers, and those of lower degree penetrate deeply. No high-degree f modes have been considered in this study.

The first inversion to be tried was with potential five-minute whole-disk data: all the modes with  $\ell = 1$  and 2 with frequencies between 2.3 and 3.8 mHz were used. It was not possible to obtain well-localized optimal kernels  $D$  at any radius, and no optimal kernel could be found with its centre below  $r = 0.6 R$ , where  $R$  is the radius of the Sun. The situation is much better for set ABS, which includes the Stanford-type data. In that case it is possible to find kernels weakly concentrated near  $r = 0.2 R$ , though not anywhere else (Figure 3). At first sight, this might seem somewhat surprising, because the amplitudes of all the splitting kernels  $K_i^*$  are greatest near the surface. It appears that localization is possible at the greatest depths because pairs of modes with different degree of like parity can be found with almost the same frequency; their kernels can be made to cancel to a considerable extent in the surface layers, where their structure depends only on  $\omega_0$ . The result is a sinuous function whose amplitude variation is substantially less than that of the raw kernels illustrated in Figure 1. There are many such pairs of kernels, and it appears that they can be combined to cancel in the region where all exist, leaving a relatively large contribution to  $D$  in the region over which the penetration radius  $L_c/\omega_0$  varies. As will be evident in subsequent inversions, when modes of higher degree are added to the data set the region over which well-localized optimal kernels can be constructed extends to larger radii.

It is now evident why the inversion of only dipole ( $\ell=1$ ) and quadrupole ( $\ell=2$ ) modes failed; amongst these there are no pairs with almost the same frequency.

The result of inverting the set ABS is shown in Figure 4. It was performed with  $\theta = 1.5 \times 10^{-3}$ . Several other values of  $\theta$  were tried, but as a result of studying the tradeoff, particularly that for set A which is illustrated in Figure 10, it was decided to adopt  $1.5 \times 10^{-3}$  as a standard value throughout, in order to save computer time. Strictly speaking the tradeoff curve is a function of the number and type of the modes that are used in the inversion. Therefore, adopting this procedure does not necessarily show off the capabilities of the inversions to best advantage. Nevertheless, since the quality of the inversions varies quite slowly with  $\theta$  in the region containing the standard value, it is unlikely that the improvements that could be achieved would be substantial. In view of the fact that even the best kernels  $D$  are only

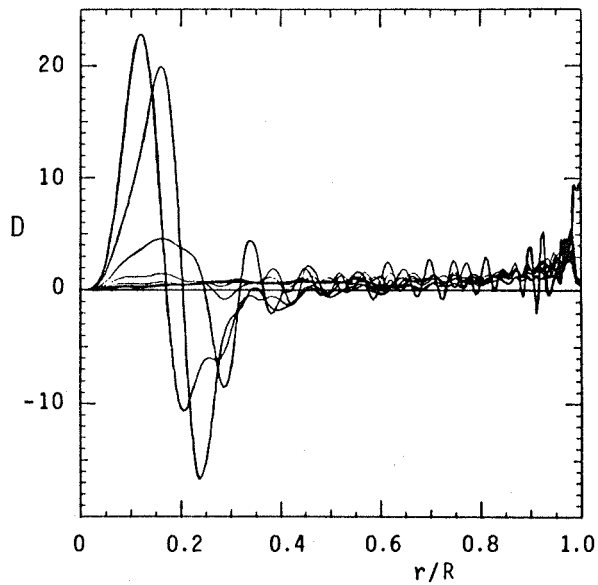


Fig. 3. A selection of optimal kernels  $D(r_0, r)$ , computed from modes ABS with  $\theta = 1.5 \times 10^{-3}$  and plotted against  $r/R$  for various values of  $r_0$ .

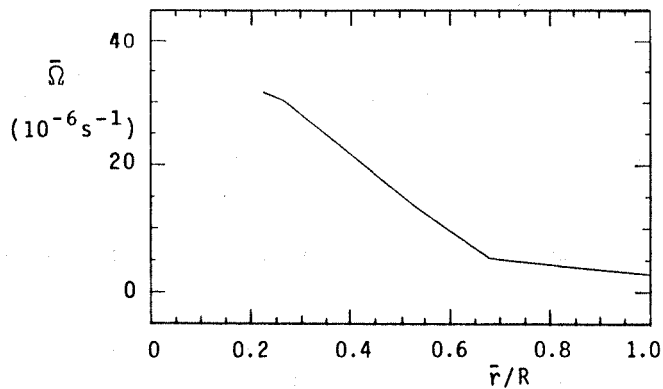


Fig. 4. Optimal averages  $\bar{\Omega}(\bar{r})$  computed from the kernels displayed in Figure 3 and joined by continuous straight lines. The estimate  $J_2[\bar{\Omega}(\bar{r})]$  of the quadrupole moment of the external gravitational potential obtained from this inversion is  $5.5 \times 10^{-6}$ .

poorly localized, and then they are not of one sign (Figure 3), one should regard the estimate  $\bar{\Omega}(\bar{r})$  in Figure 4 with considerable caution.

Inversion of the potential Duvall-Harvey set DH was much more successful. Now it is possible to find well-localized averaging kernels D everywhere except very near the centre (Figure 5). The inferred angular velocity is shown in Figure 6. It was computed with  $\theta = 1.5 \times 10^{-3}$ , at the cost of amplifying the errors by a factor  $\Lambda$  of about 4. The error magnification factor  $\Lambda$  is defined to be the ratio of the mean measure  $\epsilon$  of the error in the localized average  $\bar{\Omega}(\bar{r})$  (see Gough, 1984) averaged over all values of  $\bar{r}$ , to the mean errors in the measurements.

Figures 7 and 8 show the results of inverting the whole-disk modes WDS. Once again good resolution is possible; the optimal averaging kernels D are somewhat cleaner near the surface, and extend nearer to the centre of the Sun. It is the few g modes that are responsible for the latter. The mean error magnification factor  $\Lambda$  is 4.5.

The inversion of set A was the most thoroughly studied. With the addition of the low-frequency g modes, such as those whose kernels are shown in Figure 2, it was possible to obtain even better localization near the center of the Sun than it was with the set WDS. The error magnification factor  $\Lambda$  was also somewhat smaller for  $\theta = 1.5 \times 10^{-3}$ , though not as small as one would expect from using more than twice the number of modes.

In Figure 9 are shown inversions for three different values of  $\theta$ . In principle, the smallest value produces the highest resolution, but at the expense of a very large magnification of the errors in the data. The lowest value produces quite small errors, but the averages are over such broad ranges of  $r$  that  $\bar{\Omega}$  does not provide a good estimate of  $\Omega$ .

The points joined by the continuous lines in Figure 10 define the tradeoff curve. They are the values of the error magnification factor  $\Lambda$  plotted against the mean width  $\langle \delta \rangle$ , which is the average of  $\delta$  over the range of  $\bar{r}$  within which it was possible to centre the optimal localized averaging kernels. It was from this curve that the standard value of  $1.5 \times 10^{-3}$  for  $\theta$  was adopted, for it corresponds to a position on the bend of the curve. On either limb of the curve, the sacrifice of accuracy or resolution for a slight improvement in resolution or accuracy is much greater. All the inferred angular velocity curves  $\bar{\Omega}$  near the bend look very similar.

The final inversion was performed on set B. This is the largest data set, containing about 250 modes, and includes a wide range of p modes plus all the g modes that are contained in the other data sets. The optimal averaging kernels are shown in Figure 11; they are all quite well localized, and have centres extending to 0.04 R of the centre of the Sun. The inferred rotation curve is shown in Figure 12. The error magnification is  $\Lambda = 2.9$ .

From all the inversions of the data sets listed in Table 1 the quadrupole moment  $J_2$  of the external gravitational potential was computed from the formula

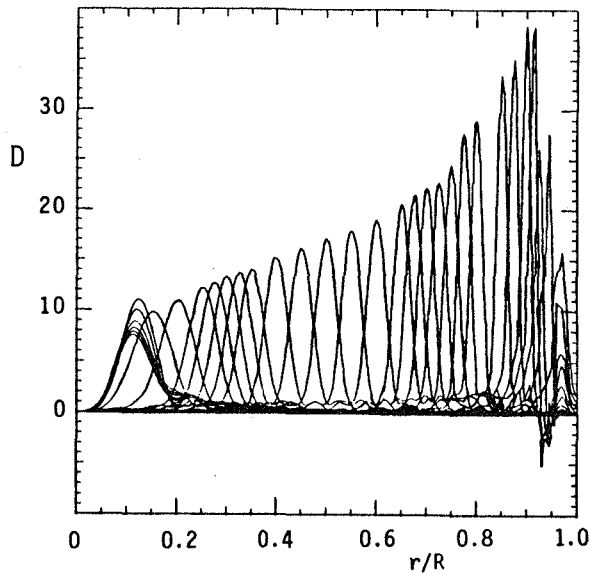


Fig. 5. A selection of optimal averaging kernels  $D$  computed from modes DH with  $\theta = 1.5 \times 10^{-3}$ .

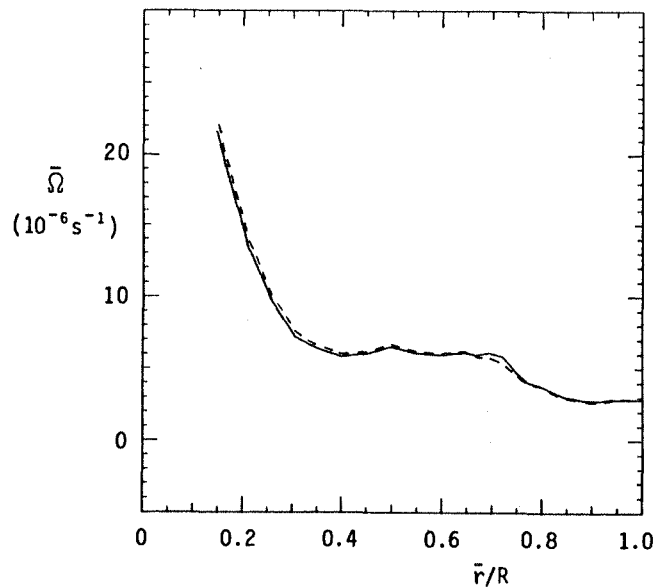


Fig. 6. Inversion of data set DH. The continuous lines join the optimal averages computed from the kernels displayed in Figure 5. The dashed lines join similar averages weighted by kernels similar to those in Figure 5 but computed from solar Model A of Christensen-Dalsgaard *et al.* (1979). The same data  $\omega_{1i}$  [computed from Model 1 of Christensen-Dalsgaard (1982)] were used in the two cases. The continuous and the dashed estimates yield  $J_2[\bar{\Omega}(\bar{r})] = 1.04 \times 10^{-6}$  and  $9.2 \times 10^{-7}$ , respectively. The error magnification factor  $\Lambda$  is 3.7 in both cases.

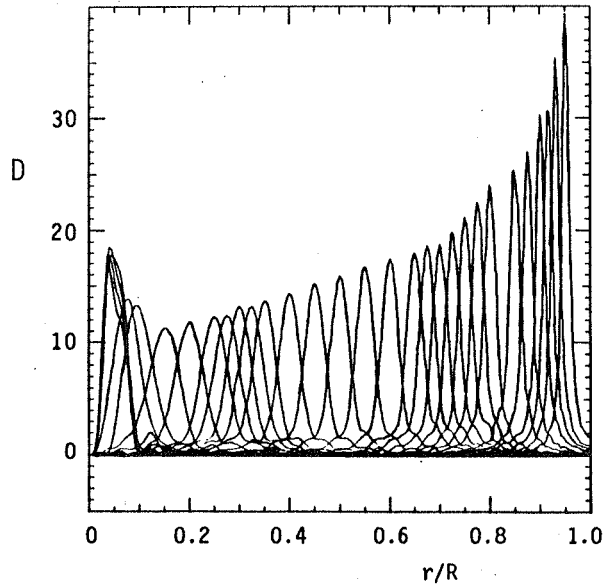


Fig. 7. Optimal averaging kernels for mode set WDS computed with  $\theta = 1.5 \times 10^{-3}$ .

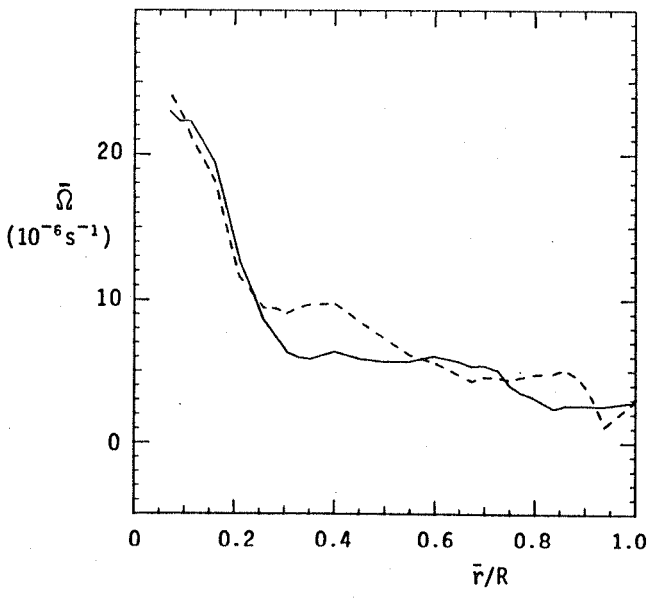


Fig. 8. Inversion of modes WDS (continuous line) using the kernels of Figure 7.  $J_2[\bar{\Omega}(\bar{r})] = 8.9 \times 10^{-7}$  and  $\Lambda = 4.5$ . The dashed line represents an inversion on Model A, which yields  $J_2[\bar{\Omega}(\bar{r})] = 1.07 \times 10^{-6}$ .

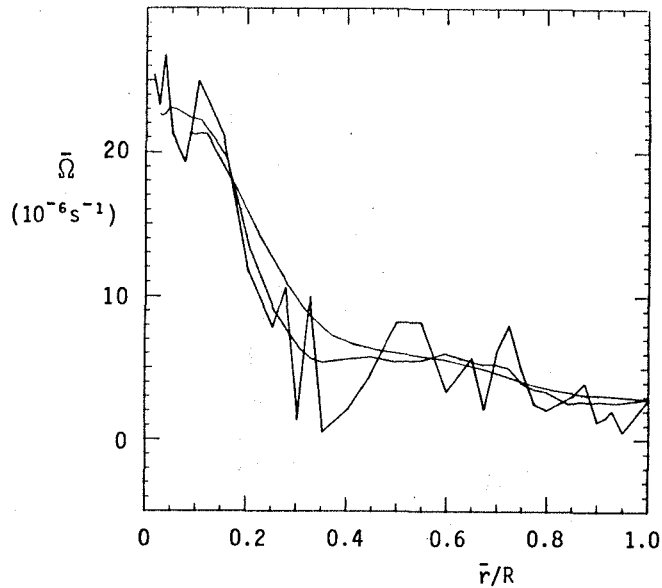


Fig. 9. Three inversions of modes A with  $\theta = 10^{-6}$ ,  $1.5 \times 10^{-3}$  and  $10^{-1}$ . Further properties of the solutions are given in Figure 10.

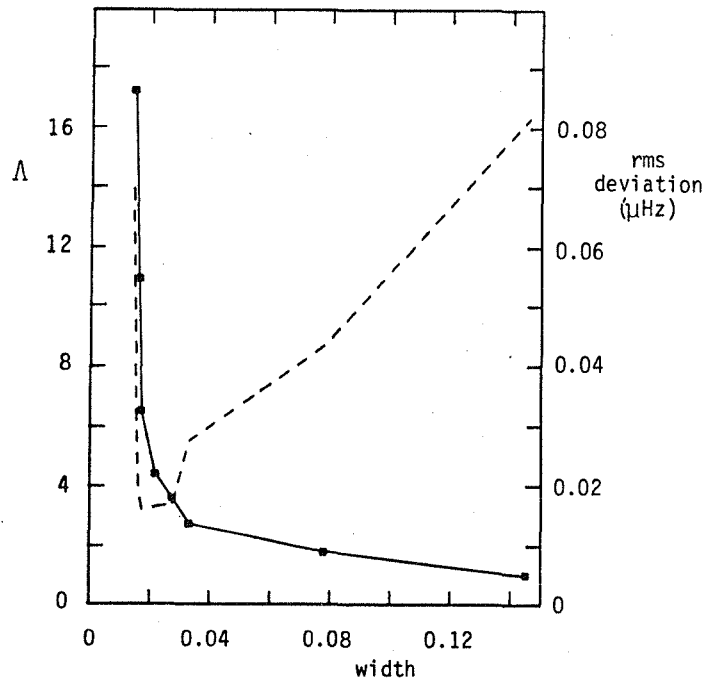


Fig. 10. The filled squares are the error magnification factors  $\Lambda$  in a selection of inversions of modes A, plotted against  $\langle \delta \rangle$ . They were computed, in order of increasing  $\delta$ , with  $\theta = 10^{-6}$ ,  $5 \times 10^{-6}$ ,  $5 \times 10^{-5}$ ,  $5 \times 10^{-4}$ ,  $1.5 \times 10^{-3}$ ,  $5 \times 10^{-3}$ ,  $2 \times 10^{-2}$ ,  $10^{-1}$ . The dashed line represents the rms deviation from the frequency splittings  $\omega_{11}$  provided for the inversion of the values computed from the averages  $\bar{\Omega}(\bar{r})$  using equation (4.5) of Gough (1984), and thus represents the accuracy to which the estimate  $\bar{\Omega}$  satisfies the data. The high values at small  $\theta$  presumably result from numerical error in the inversion procedure.

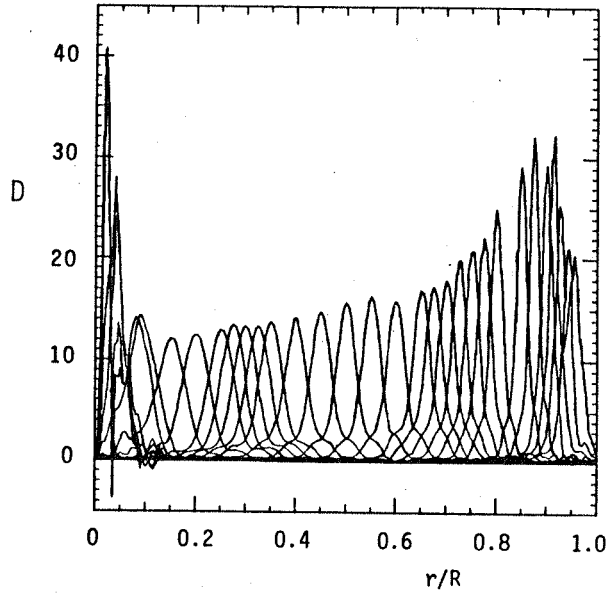


Fig. 11. Optimal averaging kernels for mode set B, computed with  $\theta = 1.5 \times 10^{-3}$ .

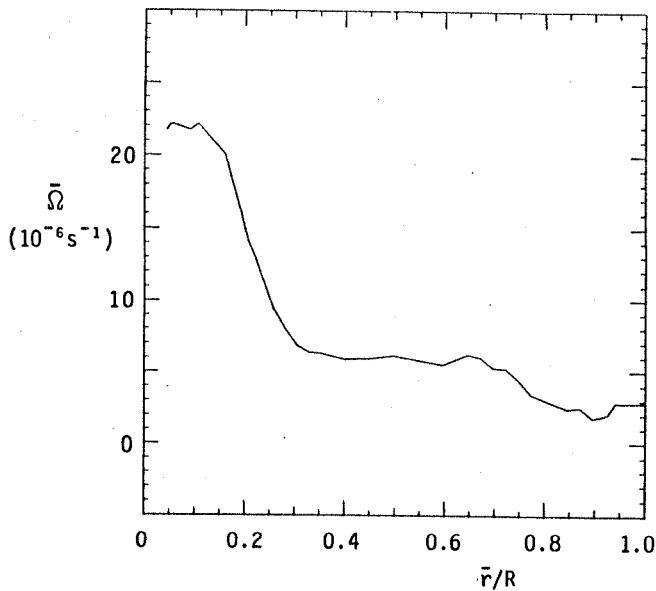


Fig. 12. Inversion of modes B using the averaging kernels of Figure 11.  $J_2[\bar{\Omega}(\bar{r})] = 9.4 \times 10^{-7}$  and  $\Lambda = 2.9$ . The rms deviation of the frequency splittings computed from  $\bar{\Omega}$  by equation (4.5) of Gough (1984) from the actual splittings provided for the inversion is  $2.6 \times 10^{-2}$   $\mu\text{Hz}$ .

$$J_2 = \int_0^R J(r) \bar{\Omega}^2(r) dr \quad , \quad (1)$$

with  $\bar{\Omega}$  represented by the localized average  $\bar{\Omega}$ . The kernel  $J(r)$  is shown in Figure 13. In order to evaluate the integral, it was necessary to extrapolate  $\bar{\Omega}$  to the origin. Little error is introduced by this procedure because the kernel  $J(r)$  is very small near the centre of the Sun.

It is interesting to consider how  $J_2$  varies along the tradeoff curve in Figure 10.  $J_2$  is a minimum ( $8.3 \times 10^{-7}$ ) at  $\theta = 5 \times 10^{-6}$ . At the only lower value of  $\theta$  considered the spurious wiggles introduced by the errors cause  $J_2$  to be larger ( $8.7 \times 10^{-7}$ ); at larger values of  $\theta$  the smoothing of  $\bar{\Omega}$  in the vicinity of the maximum of  $J$ , where the curvature of  $\bar{\Omega}$  is positive, raises  $\bar{\Omega}$  above  $\bar{\Omega}$  and so overestimates  $J_2$ , though this is offset somewhat by the smoothing near  $r = 0.7 R$  where the curvature of  $\bar{\Omega}$  is negative. The greatest estimate of  $J_2$  is  $1.06 \times 10^{-6}$ , for  $\theta = 0.1$ .

To estimate the errors in the data one might try comparing the inversions with  $\theta = 1.5 \times 10^{-3}$  in Figures 9 and 12. The rms difference between the two curves is  $0.50 \times 10^{-6} \text{ s}^{-1}$  ( $8.0 \times 10^{-2} \text{ } \mu\text{Hz}$ ), which is a little less than 5 per cent of the mean value of  $\bar{\Omega}$ . Bearing in mind that errors in the data were amplified by a factor of 2.9, one deduces that the errors in the data are about  $2.7 \times 10^{-2} \text{ } \mu\text{Hz}$  if all the differences between the two inversions result from errors in the data.

To get some idea of how important to the angular velocity inversion it is to have an accurate model of the Sun, additional inversions were carried out using the solar Model A of Christensen-Dalsgaard et al.

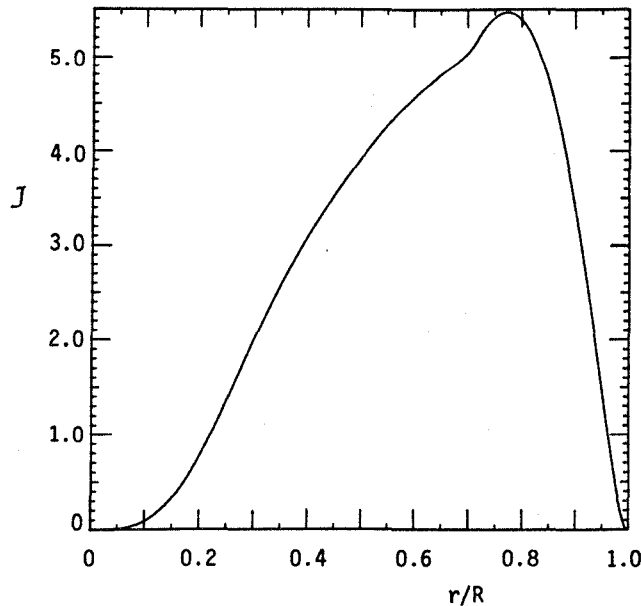


Fig. 13.  $J_2$  kernel  $J$  for Model 1, computed from the formulae (11) and (12) of Gough (1981). To compute  $J_2$  from equation (1) the values on the ordinate scale must be multiplied by  $10^{-8}$  and  $\bar{\Omega}$  must be measured in units of  $10^{-6} \text{ s}^{-1}$ .



(1979) to compute the splitting kernels  $K_1^*$ . The results of inverting the splitting frequencies of sets DH and WDS (which were computed from Model 1) are shown in Figures 6 and 8. It is interesting that when modes with a wide range of  $\ell$  are used, which penetrate to a variety of different depths, the result is less sensitive to the equilibrium solar model than when the whole-disk modes are used. Roughly speaking, subtraction of modes of different but similar degree and almost identical frequencies tends to give local information situated between the penetration depths of the two modes. Cancellation of the eigenfunctions at larger radii takes place to a similar degree whatever the solar model. On the other hand, the low-degree modes of WDS all sample essentially the whole range of  $r$ . Inverting the data is more like computing the coefficients of a Fourier expansion of a function  $f$ , each of which depends on the structure of  $f$  over the entire domain. Under these circumstances the errors introduced by misrepresenting the equilibrium model are greater.

Finally, the value of  $J_2$ : The most accurate inversion is probably the last, in Figure 12. With the error estimate inferred above, one obtains  $J_2 = 9.4 \pm 0.9 \times 10^{-7}$ .

### THIRD STAGE: COMPARISON WITH THE ORIGINAL ARTIFICIAL ROTATION

The angular velocity used to generate the splitting data is shown in Figure 14. It is compared with the inversion that was judged the best. For this,  $\bar{\Omega}$  is plotted with vertical bars to represent the estimated error in its value arising from the errors in the data. The horizontal bars extend to  $\pm\delta/2$ , indicating the width of the averaging kernel. The random errors that were added to the computed frequency splitting had standard deviation  $\sigma = 0.01 \mu\text{Hz}$ . The value of  $J_2$  computed from the actual angular velocity is  $8.9 \times 10^{-7}$ .

### CONCLUSION

It can be seen from Figure 14 that except near the very centre of the Sun the estimate of  $\Omega(r)$  made by the observer (Author B) is consistent with the true angular velocity with which the solar model had been endowed. If the estimated error in the optimized averages  $\bar{\Omega}$  is assumed to arise solely from errors in the data, the latter are overestimated by a factor of nearly 3. Therefore, we conclude that much of the error in  $\bar{\Omega}$  is a product of inadequate resolution, and that a substantially improved inversion could probably have been achieved had more modes been used. The gravitational quadrupole moment  $J_2$  of the solar model was correctly predicted by the inversion, within the limits set by the uncertainty in the latter.

It appears, therefore, that a good inversion of rotational splitting could be carried out if data from modes with either a wide range of  $\ell$ , such as those that have been observed by Harvey and Duvall (1984), or a wide range of  $n$ , such as might be measured from whole-disk observations

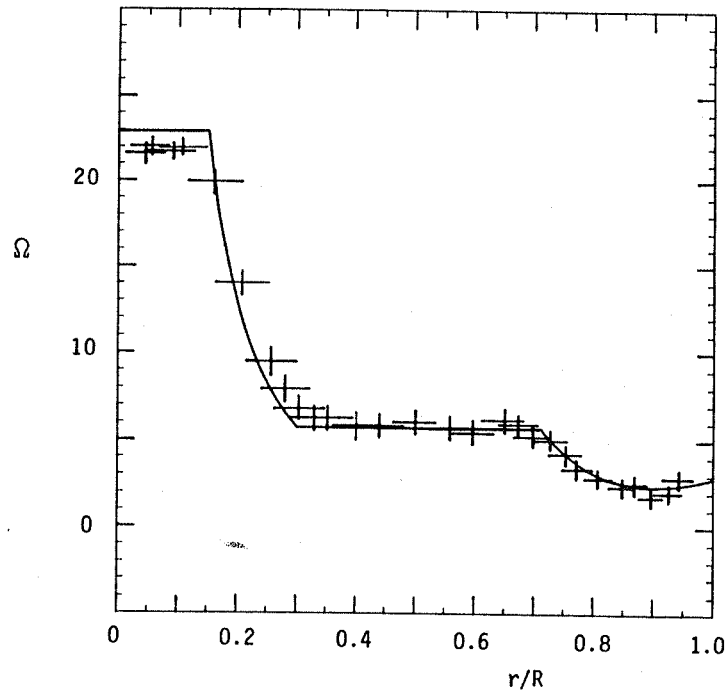


Fig. 14. The continuous line is the angular velocity  $\Omega(r)$  from which the splitting frequencies  $\omega_{11}$  used in the inversions were computed. The crosses are centred at a selection of the averages  $\bar{\Omega}$  obtained from inverting modes B and shown in Figure 12 (a few were omitted to avoid cluttering the diagram). The vertical components of the crosses represent the estimate of the rms error in  $\bar{\Omega}$  based on the error estimate made in Stage Two, and the horizontal components are of length  $\delta/R$ .

from space, could be obtained. The result of the former, if the splitting for all values of  $m$  could be measured, would give us not only information about the radial dependence of  $\Omega$ , but also about the latitudinal dependence (cf. Gough, 1982). The latter would provide only broad averages over latitude. Whole-disk rotational splitting data from five-minute dipole and quadrupole oscillations would be quite inadequate to infer  $\Omega(r)$ , though some idea of the general form of  $\Omega(r)$  could be obtained if these data were supplemented with appropriate measurements of the type that might be obtained from Stanford. The addition of identified low-degree  $g$  modes, of the type reported by Scherrer and Delache (1984), would improve the inversion considerably, especially near the centre of the Sun.

One might regard the inversions carried out here to present an overoptimistic view of what might be deduced from real solar data, because the errors are lower than those one might anticipate having to contend with from imminent ground-based observations. However, on the longer term, data from space or from a ground-based network of observatories would be expected to provide splitting from many more modes than have been considered in this exercise. Not only would this increase the cancellation of the errors, but it would also improve the

representation  $\bar{\Omega}$  (or a representation by a spectral expansion) by virtue of achieving higher resolving power. One might also argue that we have overstated our case by using the same kernels  $K_i^*$  for the inversion as we did to compute the artificial splitting data when offering our best estimate of  $\Omega$ . This is unlikely to be an important issue. As is shown in Figure 6, the inferred angular velocity is insensitive to the solar model if modes with a wide range of  $\ell$  are available. And in any case, by the time copious splitting data  $\omega_{ij}$  are available there should be sufficient information about the stratification of the Sun obtained from inverting the frequencies of axisymmetrical modes to reduce the errors in  $K_i^*$  well below the errors in  $\omega_{ij}$ .

#### REFERENCES

- Backus, G. and Gilbert, F. 1970: Phil. Trans., 266A, 123-192  
 Christensen-Dalsgaard, J. 1982: Mon. Not. R. astr. Soc., 199, 735-761  
 Christensen-Dalsgaard, J. and Frandsen, S. 1983: Solar Phys., 82, 469-486  
 Christensen-Dalsgaard, J., Gough, D. O. and Morgan, J. G. 1979: Astron. Astrophys., 73, 121-128; 79, 260  
 Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. and Roca Cortez, T. 1980: Astron. Astrophys., 91, L9-L10  
 Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. and Roca Cortez, T. 1981: Nature, 293, 443-445  
 Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B. and Roca Cortez, T. 1982: Solar Phys., 74, 51-57  
 Duvall, T. L., Jr and Harvey, J. W. 1983: Nature, 302, 24-27  
 Grec, G., Fossat, E. and Pomerantz, M. 1980: Nature, 288, 541-544  
 Grec, G., Fossat, E. and Pomerantz, M. 1983: Solar Phys., 82, 55-66  
 Goldreich, P. and Keeley, D. A. 1977: Astrophys. J., 212, 243-251  
 Gough, D. O. 1980: In Nonradial and Nonlinear Stellar Pulsations (ed. H. A. Hill and W. A. Dziembowski, Springer, Heidelberg), pp. 273-299  
 Gough, D. O. 1981: Mon. Not. R. astr. Soc., 196, 731-745  
 Gough, D. O. 1982: Nature, 298, 334-339  
 Gough, D. O. 1984: these proceedings  
 Harvey, J. W. and Duvall, T. L., Jr 1984: these proceedings  
 Hill, H. A., Bos, R. J. and Goode, P. R. 1982: Phys. Rev. Lett., 49, 1794-1797  
 Scherrer, P. and Delache, P. 1984: Mem. Soc. astr. Italiana, in press  
 Scherrer, P. H., Wilcox, J. M., Christensen-Dalsgaard, J. and Gough, D. O. 1982: Nature, 297, 312-313  
 Scherrer, P. H., Wilcox, J. M., Christensen-Dalsgaard, J. and Gough, D. O. 1983: Solar Phys., 82, 75-87



## SENSITIVITY OF INFERRED SUBPHOTOSPHERIC VELOCITY FIELD TO MODE SELECTION, ANALYSIS TECHNIQUE AND NOISE

### **FRANK HILL**

National Solar Observatory, Sunspot, NM 88349, U.S.A., and  
Joint Institute for Laboratory Astrophysics, University of  
Colorado

### **DOUGLAS GOUGH**

Institute of Astronomy, and Department of Applied Mathematics  
and Theoretical Physics, University of Cambridge, Cambridge,  
CB3 0HA, England, and Joint Institute for Laboratory  
Astrophysics, University of Colorado

### **JURI TOOMRE**

Department of Astrophysical, Planetary and Atmospheric  
Sciences, and Joint Institute for Laboratory Astrophysics,  
University of Colorado, Boulder, CO 80309, U.S.A.

### **ABSTRACT**

The horizontal velocity immediately below the photosphere can be inferred from observations of high-degree solar oscillations by an optimal-averaging inversion technique. We investigate the sensitivity of the results to various details of both the inversion and the determination of the frequencies. The results are shown to be quite stable to the choice of most parameters, suggesting that this procedure produces reliable estimates of the subsurface velocity.

### **1. INTRODUCTION**

Helioseismology has now developed to the point where we are able to infer the physical conditions of the solar interior from observations of the frequencies of the Sun's oscillations. Owing to its comparatively simple effect on the frequencies, the quantity most readily obtained is horizontal velocity. This velocity is primarily differential rotation, though it no doubt contains a contribution from convection cells. Deubner, Ulrich and Rhodes (1979) have attempted to determine the solar rotation from five-minute oscillation frequencies, though their method did not permit them to determine the depth dependence. A similar analysis applied to later observations failed to reproduce the earlier results (Rhodes, Harvey and Duvall 1983).

The prospect of detecting giant cells from their effects on the oscillations is exciting, particularly because earlier searches for surface Doppler velocities with a large spatial scale have failed. This suggests that the photospheric velocity amplitude of giant cells is below the  $10 \text{ ms}^{-1}$  sensitivity of the measurements (LaBonte, Howard and Gilman 1981). However, a theoretical model of compressible convection predicts that the horizontal velocity increases with depth (Latour, Toomre and Zahn 1983). We thus might expect to find subsurface flows with amplitudes substantially greater than the limits set by the photospheric observations.

We have sought to infer the horizontal velocity as a function of depth from the positions of the ridges in the  $k$ - $\omega$  diagram of high-degree ( $\ell > 100$ ) five-minute oscillations. Here  $k$  is the horizontal wavenumber and  $\omega$  is the temporal frequency. Figure 1 displays a contour plot of one of our observed  $k$ - $\omega$  diagrams, showing the ridges corresponding to the  $f$ , or fundamental, mode and the  $p_1$  through  $p_8$  modes. We have previously reported (Hill, Toomre, and November 1982, 1983) variations in the ridge positions from day to day that were possibly associated with the passage of giant cells across the field of view (e.g. Gough and Toomre 1983).

The possibility of inferring subsurface velocities stems from the fact that wave patterns are advected by a horizontal flow (e.g. Rhodes, Deubner and Ulrich 1979). In the case of high-degree modes, the wavelength of the oscillations is much smaller than the horizontal scale of the flow pattern of interest, namely the giant cells. Then the apparent frequency of a mode with a given  $(k, \omega)$  is modified by an amount  $\delta\omega = k \bar{U}$ . Provided the variation at all detectable depths of the horizontal component of the subphotospheric flow velocity  $U$  across the field of view and throughout the observing interval can be ignored, the advection velocity  $\bar{U}$  of the wave pattern is the vertical average of the equatorial component of  $U$  weighted by the energy density of the mode (Gough 1978).

We have applied an optimal averaging inversion procedure to perturbations of the ridge position that we believe are caused by large-scale horizontal flows. The inversion procedure is suggested by the work of Backus and Gilbert (1968, 1970) and has been most fully developed in the field of geophysics. In our case, the frequency shift  $\delta\omega$  induced by the flow  $U$  is

$$\delta\omega_i = k_i \bar{U} = k_i \int_{\zeta_1}^{\zeta_2} A_i U H d\zeta, \quad (1)$$

where  $\zeta = \log_{10} p$  is used as the depth coordinate ( $p$  is the pressure in the equilibrium model),  $H$  is the pressure scale height and  $A_i$  is proportional to the kinetic energy density of the mode designated by the subscript  $i$ . The constant of proportionality is chosen to render  $H A_i$  unimodular; the limits  $\zeta_1$  and  $\zeta_2$  bound the interval within which  $H A_i$  differs significantly from zero. The inversion procedure involves the construction of linear combinations of equation (1), yielding

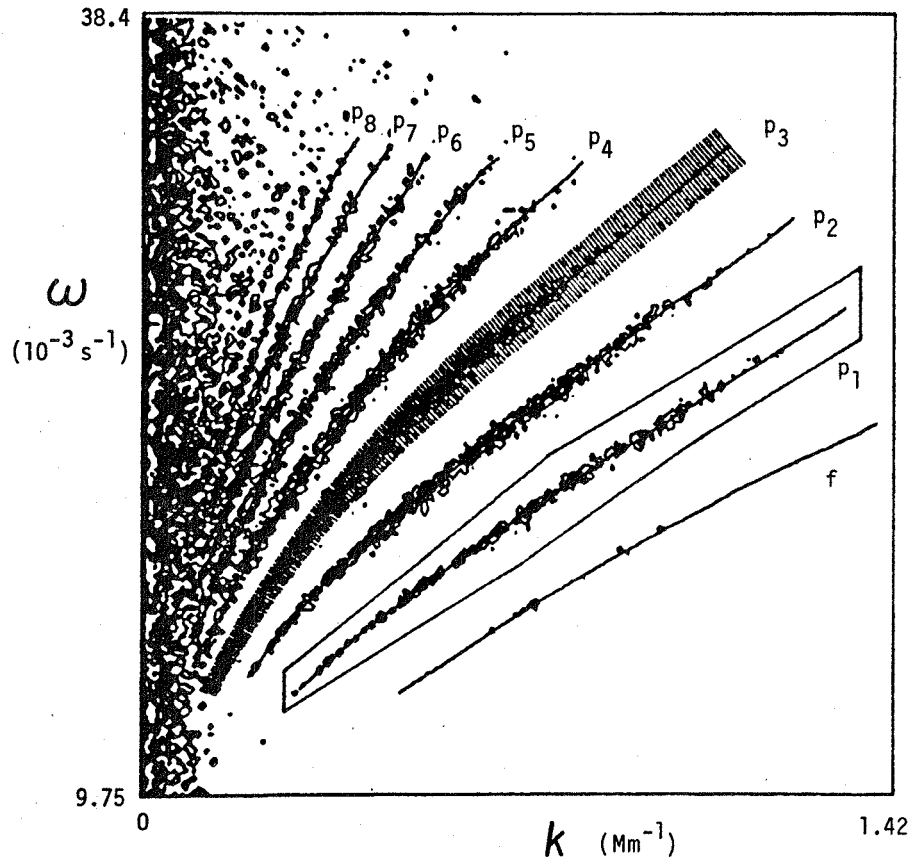


Figure 1. Contours of constant-power in the  $k$ - $\omega$  plane obtained on day 82 (23 March 1981). The fits resulting from the iterative ridge-fitting procedure are overlaid on the diagram. One of the initial boxes used to isolate the ridges is drawn around the  $p_1$  ridge. The cross-sections used to determine the ridge positions are indicated on the  $p_3$  ridge. Our ridge classification ignores the possible existence of chromospheric modes.

$$\int_{\zeta_1}^{\zeta_2} \sum_i \alpha_i(\zeta) A_i(\zeta') U(\zeta') H d\zeta' \equiv \int_{\zeta_1}^{\zeta_2} D(\zeta; \zeta') U(\zeta') d\zeta'$$

$$\equiv \tilde{U}(\zeta) = \sum_i \alpha_i k_i^{-1} \delta\omega_i, \quad (2)$$

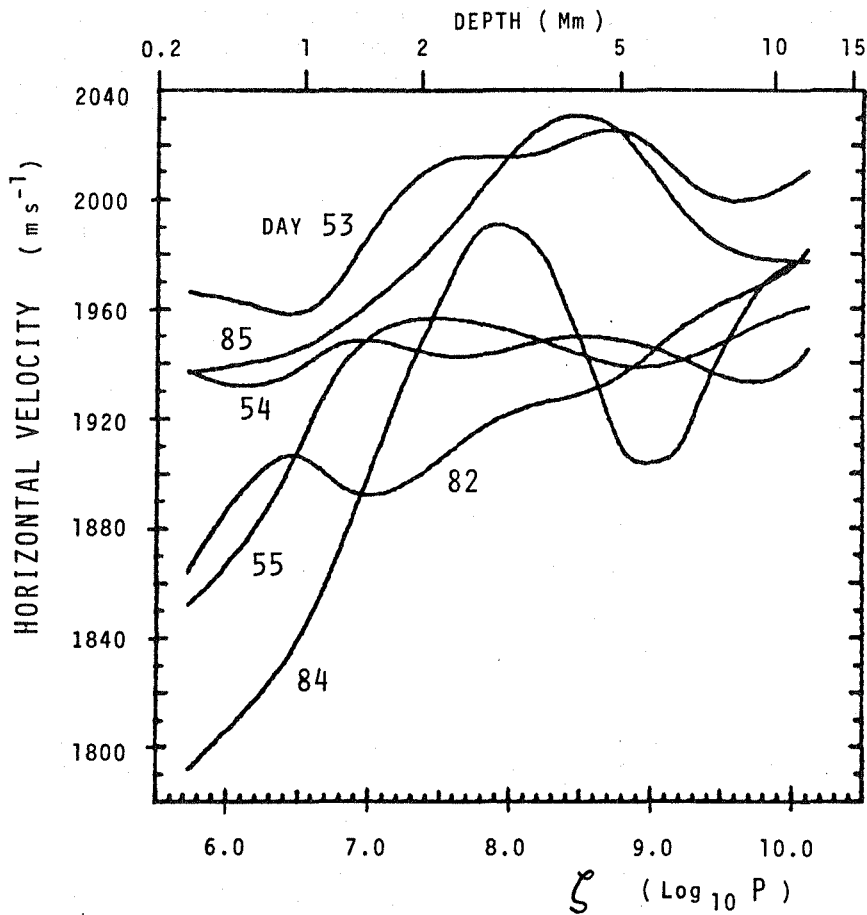
The coefficients  $\alpha_i$  are chosen to make  $D(\zeta, \zeta')$  resemble a Dirac delta function  $\delta(\zeta - \zeta')$  as closely as possible. Then  $\tilde{U}(\zeta)$  estimates  $U(\zeta)$ . The technique of Backus and Gilbert (1970), which takes account of the errors in the data, was used to calculate the coefficients  $\alpha_i$ . More details of inversion procedures applied to similar data can be found in Hill, Gough and Toomre (1984), Gough (1984), and Christensen-Dalsgaard and Gough (1984).

The inversion was carried out on  $k$ - $\omega$  diagrams computed from observations of Doppler velocities obtained at Sacramento Peak Observatory (SPO). Data were obtained from the Fe I 5434.5 Å and Mg I 5172.7 Å spectral lines using the diode array situated at the exit of the echelle spectrograph of the vacuum tower telescope. Sequential two-dimensional intensity images in the red and blue wings of both lines were produced by spatially scanning the Sun across the slit of the spectrograph. The scan produced a 256"  $\times$  1024" image centered on the solar equator with a nominal resolution of 2". Scans were repeated every 65 or 70 s for observing intervals of 7 to 11 hours.

The data reduction began with the determination of the line-of-sight velocity in the manner described by November et al. (1979). The resulting velocities were averaged perpendicular to the equator to filter out all waves except those with a horizontal wavenumber vector oriented nearly parallel to the equator. The data were projected onto an array with a uniform spacing in longitude of 0.1183°, corresponding to a resolution in  $k$  of  $8.54 \times 10^{-3} \text{ Mm}^{-1}$ . The time series was extended with zeros to obtain a constant resolution of  $1.5 \times 10^{-4} \text{ s}^{-1}$  in  $\omega$ . The resulting data array was then Fourier transformed to produce a  $k$ - $\omega$  diagram such as that in Figure 1. We have considered a total of six different  $k$ - $\omega$  diagrams obtained in February and March of 1981. We shall identify the data by the number of the day in the year when they were obtained. Figure 1 shows the diagram for day 82, which is 23 March 1981.

We have inverted the data to produce an estimate of the horizontal velocity in the 15 Mm immediately below the photosphere. The results are shown in Figure 2, which displays the results for the six different days. There is a general tendency for the horizontal velocity to increase with depth, and there is a variation of the order of  $100 \text{ ms}^{-1}$  from day to day in the curves. We suggest that these curves possibly reflect an increasing rotational velocity with depth on which is superimposed a large-scale convective flow of about  $100 \text{ ms}^{-1}$ . Further discussion of the interpretation of these curves can be found in Hill, Gough and Toomre (1984). What we discuss in this paper is the sensitivity of the results of the inversion to variations of certain details of the data analysis.





**Figure 2.** Horizontal velocity as a function of depth  $\zeta$  ( $\equiv \log_{10} p$ ) inferred from the inversion of data obtained on 6 days in 1981. The curves are identified by the number of the day of the year on which the observations were made. There is a general increase of velocity with depth, with day-to-day variation on the order of  $100 \text{ ms}^{-1}$ . The latter is probably the result of the changing position of giant cells, and is superimposed on a differential rotation velocity that increases with depth.

## 2. DETERMINATION OF THE FREQUENCIES

Since the individual modes cannot be resolved by these observations, we have estimated the frequencies by a ridge-fitting technique. It appears that this method is useful in overcoming some of the effects of mode beating and atmospheric seeing that are a source of noise. From the size of our resolution bins and the distribution of the solar modes in the  $k$ - $\omega$  diagram, we estimate that we have about 30 modes in each bin. The beating of these modes produces a jagged mountain-chain appearance to the ridges which changes markedly from day to day. Mode beating can actually place power outside the bin in which the responsible modes reside, though the extent to which this occurs remains to be investigated. In addition, image motion produced by atmospheric seeing can alter the distribution of power along a ridge (Hill 1984). The net result of these processes is to impart a small-scale variation to the location of the ridges which we must try to eliminate in our search for frequency changes due to solar effects. We have thus developed a technique of ridge finding which involves a considerable amount of smoothing.

The ridge-fitting procedure begins with the drawing of a polygonal box around each of the 18 ridges in the  $+\omega$  and  $-\omega$  quadrants of the  $k$ - $\omega$  diagram having  $k > 0$ , using the cursor of an interactive graphics device to choose the vertices. The boxes define the limits of slices in the  $k$ - $\omega$  diagram that are used to calculate initial estimates of the median positions of the ridges. One of these boxes, drawn around the  $p_1$  ridge, is shown in Figure 1. The orientations of the initial slices are vertical (constant  $k$ ) for the  $f$  ridge and the  $p_1$  through  $p_3$  ridges, and horizontal (constant  $\omega$ ) for the  $p_4$  through  $p_8$  ridges. Different orientations are chosen because the higher-order ridges have steeper slopes.

The data along each slice are interpolated onto a grid of 0.1 pixel using a cubic spline. The median of the smoothed ridge cross-section, defined to be the point that divides the integral of the power in half, is computed. The loci of the medians in the  $k$ - $\omega$  plane are then smoothed with a running mean and fitted with a variable-knot cubic spline. This type of cubic spline allows the specification of the number and the initial positions of knots at which the spline conditions are applied. These positions are then varied to minimize the rms deviation between the fit and the raw medians. In all cases, the initial knots were approximately evenly spaced along the ridges.

Our original data analysis ended at this point, and used the resulting fit computed with 7 knots to provide the frequencies of the modes. We have found, however, that performing the inversion on the frequencies thus obtained gave inconsistent results when we considered different sets of modes that sample similar depths. Further investigation showed that the shape of the boxes drawn about the ridges markedly influenced the frequencies. In addition, the loci of medians computed from cross-sections parallel to the  $k$  and  $\omega$  axes are different. This

arises in part from the curvature of those loci. Thus we have turned to an iterative scheme of ridge finding that calculates cross-sections perpendicular to the ridges. We have also attempted to minimize the influence of noise in the power spectrum by using the positions of adjacent ridges to choose the edges of the boxes.

The orthogonal ridge fitting begins from an initial estimate obtained from the procedure described above, using 3 knots. This is used to estimate the positions of the adjacent ridges which, where necessary, are extrapolated using the functional form deduced by Duvall (1982). At any given  $k$  the direction perpendicular to the initial fit is then calculated from the cubic spline coefficients, and the intersection between this line and the adjacent ridges is determined. The values of  $k$  at which these perpendicular slices are calculated are evenly spaced in arc length along the ridge.

Next, a small region centered on the ridge at the chosen  $k$  is used to provide the coefficients of a two-dimensional cubic spline interpolation. These coefficients determine what we call a slice: the interpolated cross-section of the power along the direction perpendicular to the ridge, with resolution of 0.1 pixel. The slice is then truncated on both sides of the ridge, the points of truncation being determined by a parameter which we call the cut factor; it is defined as a fraction of the distance along the slice to the estimated adjacent ridge. The directions and lengths of the slices for a cut factor of 0.3 are shown superimposed on the  $p_3$  ridge in Figure 1.

The medians of the splines along the slices (which we call the raw medians) are determined, and their loci are smoothed, using a 3-knot cubic spline. The latter we refer to as a ridge fit. This provides a first iterate of the position of the ridge. The procedure is repeated once more using the iterate to approximate the adjacent ridge positions and to provide the directions of the cross-sections. The final fit is performed with 7 knots, in order to include intermediate-scale structure along the ridge. Five points are removed from each end of the ridge before the final fit to remove end effects introduced by the running mean.

### 3. SENSITIVITY TO ANALYSIS TECHNIQUE

We have varied some of the parameters that enter into our rather complicated ridge-finding procedure to assess the sensitivity of the inversion to the somewhat arbitrary details of the analysis. The parameters that we have varied are: the number of knots in the final fit, the number of knots in the intermediate fits, the number of points in the running mean used to smooth the final fit, the number of points in the running mean used to smooth the intermediate fits, and the cut factor. We have also computed medians using slices truncated symmetrically on both sides of the ridge.

The results are shown in Figures 4 through 9. All of the figures present the horizontal velocity as a function of depth inferred from the inversion of the data obtained on day 82. The curve that we feel is most reliable was produced from the orthogonal ridge-fitting procedure using a 15-point running mean to smooth both the final and intermediate medians: 7 knots for the final fit, 3 knots for the intermediate fits, and a cut factor of 0.3 to each ridge. This curve can be seen in Figure 2.

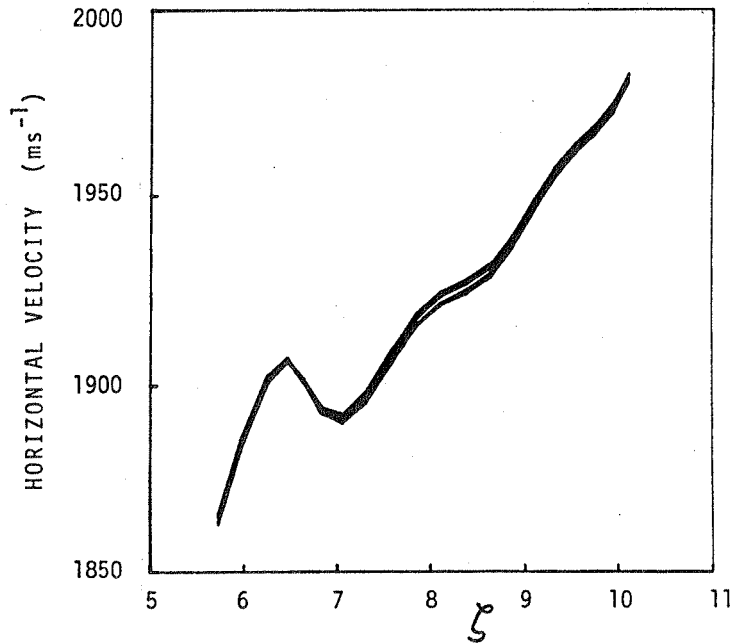
In Figure 3 we show the results of varying the number of points in the running mean that is used to smooth the final set of medians. A total of 10 different uniformly weighted means were used, ranging from 1 (no running mean at all) to 19 points, with every intervening value being an odd number. All 10 curves are plotted in Figure 3, which shows that the procedure is quite stable to variations of this parameter; the greatest discrepancy between the curves is about  $5 \text{ ms}^{-1}$ .

Figure 4 shows the result of varying the extent of the running mean applied to the first iterates on the medians; the same selection of widths was used as for the final running mean discussed above. The variation, although still small, is more substantial here, being at most  $12 \text{ ms}^{-1}$ . This possibly reflects the effect of small differences in the direction of the line along which the cross-section is computed.

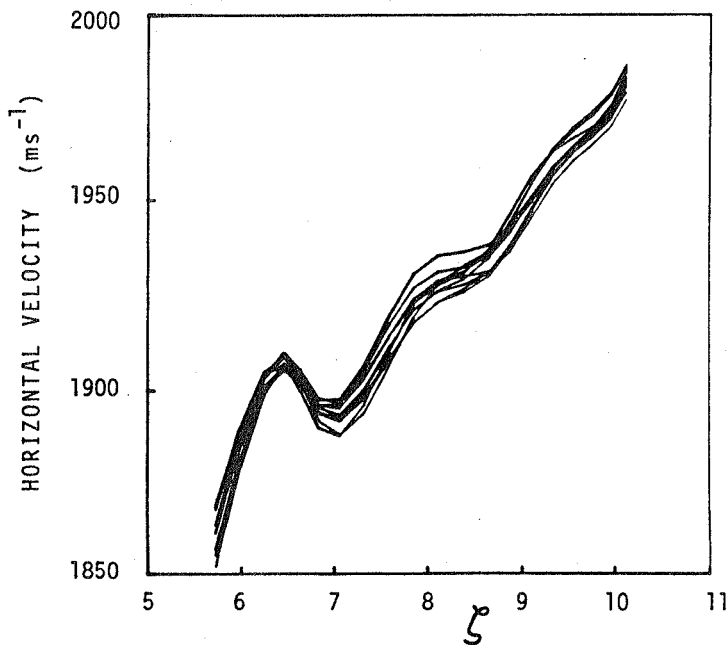
Figure 5 shows the effect of varying the number of knots used to define the final fit. A total of six cases was computed, using 3, 5, 7, 10, 15, and 20 knots. Figure 5 shows that the resulting velocity curve is stable to the final number of knots as long as it is above 3. Though the data analysis is designed to suppress small-scale variation along the ridge, we must not obliterate everything. We expect to be able to extract intermediate scales from the noise. It appears that a mere three knots are insufficient for this purpose. The maximum variation for the cases where the number of knots is 5 or greater is about  $5 \text{ ms}^{-1}$ . On the other hand, the procedure proved to be incapable of successfully finding the complete ridge when an intermediate fit of greater than 3 knots was used. This arises from changes in the ridge slicing direction. When many knots are used to compute the intermediate fit, the small-scale structure results in a rapidly varying slicing direction and renders the procedure unstable.

Figure 6 compares our standard asymmetrical truncation of the slices with a symmetrical truncation. In the latter case the two truncation points were the same distance from the median, and were a constant factor of the distance to the nearest ridge. Thus noise is sampled differently. The distinction between the two methods is not important, however, as the maximum variation between the two curves is only  $5 \text{ ms}^{-1}$ .

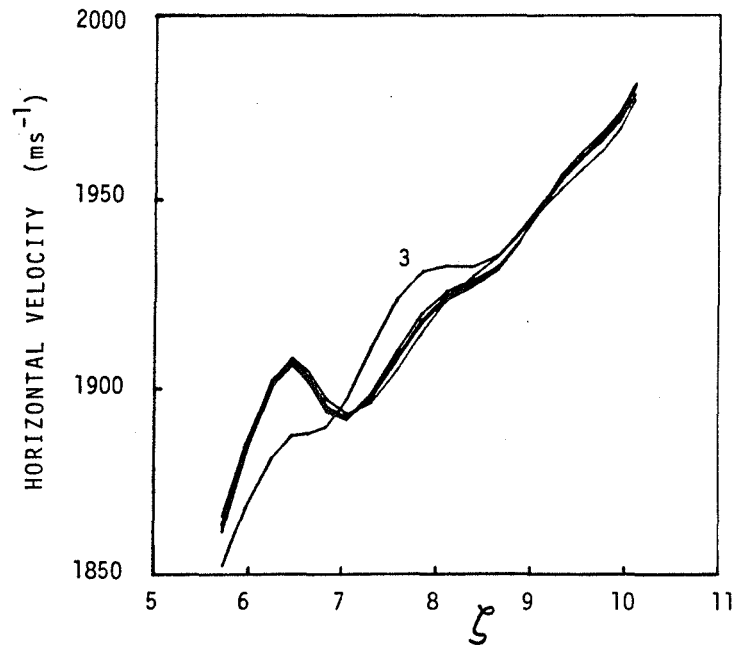
Figure 7 illustrates the result of varying the cut factor. Values of 0.10, 0.15, 0.20, 0.25, 0.30, and 0.35 have been tried. The results indicate that the procedure is rather sensitive to this parameter, for



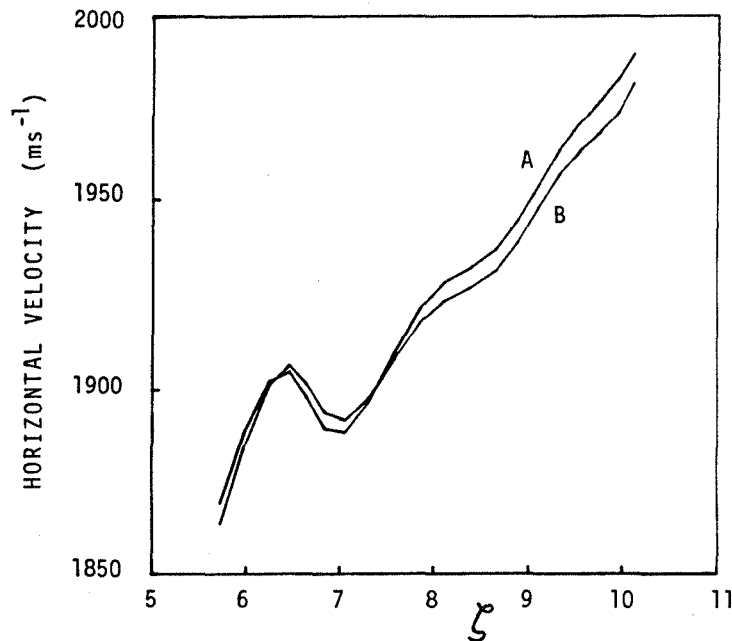
**Figure 3.** The horizontal velocity as a function of depth  $\zeta$  on day 82 calculated using 10 different widths for the final running mean. The maximum variation between the curves is  $5 \text{ ms}^{-1}$ , indicating that the inversion procedure is stable to variations of this parameter.



**Figure 4.** As Figure 3, but for 10 different widths of the intermediate running mean. The variation is greater here being about  $12 \text{ ms}^{-1}$ . This is due to the effect of small changes in the direction along which the cross-section is computed.



**Figure 5.** As for Figure 3, but using different numbers of knots in the final fit of the medians. The curve labeled 3 was obtained from 3 knots, and no doubt fails to reflect the structure along the ridge. The results are stable provided the number of knots is no less than 5.



**Figure 6.** As for Figure 3, but comparing the effects of symmetrically truncated slices whose lengths are determined by the distance to the nearest ridge (curve A) and asymmetrical truncations determined by the positions of both ridges (curve B).

the variation of the inferred velocity can be as much as  $50 \text{ ms}^{-1}$ . The sensitivity is least for cut factors in the narrow range  $0.25 - 0.35$ , where the maximum variation is  $7 \text{ ms}^{-1}$ . Cut factors smaller than  $0.25$  result in the loss of the edges of the ridges and do not provide an accurate median. The ridge finding procedure is incapable of finding the complete ridges when the cut factor is greater than  $0.35$ , because larger cross sections encompass power from adjacent ridges in some locations, which pulls the fit away from the desired ridge in the iteration.

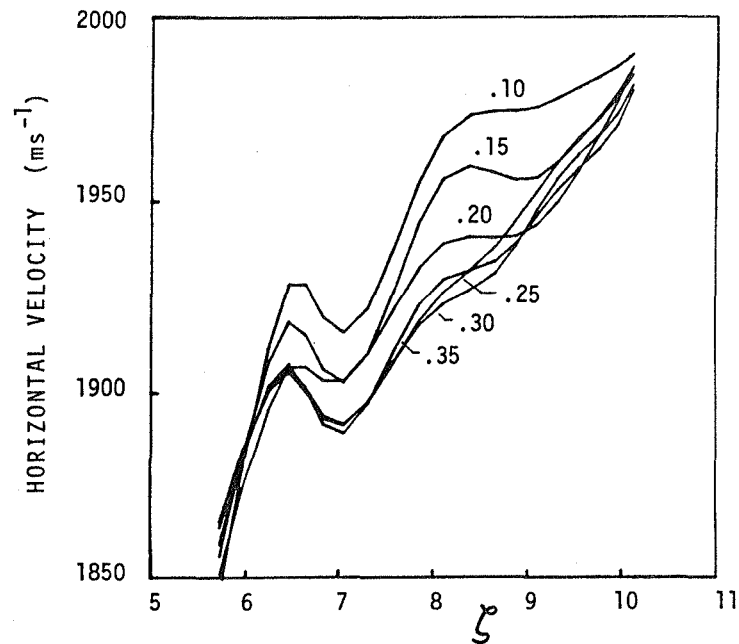
#### 4. SENSITIVITY TO MODE SELECTION AND NOISE

The inversion technique results in the construction of a set of optimal averaging kernels  $D(\zeta, \zeta')$ . As is evident from equation (2), each optimal kernel provides a weight function to average the flow velocity  $U$ , the result of which is a weighted average of the data. Because the original kernels  $A_i$  effectively span a space of dimension less than the number of data, there can be redundancy in the data, and the combinations that yield  $D$  reduce the effects of noise. However, there is a tradeoff between resolution and error. A narrower kernel  $D$  requires coefficients  $\alpha_i$  with greater magnitudes, resulting in a greater contribution from uncorrelated noise to the inferred velocity. Thus it is expedient to sacrifice some resolution.

The inversion technique takes errors into account by means of a parameter  $\theta$  whose value can vary from  $0$  to  $\pi/2$  (Backus and Gilbert 1970; see also Gough 1984). When  $\theta = 0$ , errors in the data are ignored and the widths of the optimal kernels are minimized to provide the highest resolution. Random errors are greatly magnified, and yield an unrealistically rapidly varying velocity. On the other hand, when  $\theta = \pi/2$ , the kernels  $A_i$  are ignored entirely when determining  $\alpha_i$ , and the velocity curve that is produced has unacceptably poor resolution. An intermediate value of  $\theta$  must be chosen by performing the inversion for a number of different choices for  $\theta$  and selecting a value at which the shape of the velocity curve is stable. If the inversion is nowhere stable, the data are too heavily contaminated by error to permit an inference of the flow velocity, unless one is prepared to assume some constraint on  $U$ .

The procedure is illustrated in Figure 8, which shows the result of the inversion with  $\theta = 0.001, 0.01, 0.1, 0.5, 0.7, \text{ and } 1.0$ . The first three values produce widely varying velocities characteristic of inversions that ignore errors. The curves stabilize at  $\theta = 0.5$ , and the smaller-scale structure begins to be smoothed out as  $\theta$  approaches  $\pi/2$ . This can be seen more clearly in Figure 8b, which shows the curves for  $\theta = 0.5, 0.7, \text{ and } 1.0$  on an expanded scale. Figure 8b also shows that the depths at which the optimal kernels are centered changes as  $\theta$  is varied. We have chosen  $\theta = 0.5$ .

The effectiveness of the inversion procedure in reducing errors is shown by comparing the noise in the raw data with the errors in the inferred velocity. We have defined the noise in the data to be the rms



**Figure 7.** As for Figure 3, but comparing the results from different lengths of the cross-section. The curves are labeled with the cut factor used to compute them. The cut factor is the fraction of the distance to the next ridge that is covered by the cross-section. The results vary considerably, except in the narrow range 0.25 - 0.35. We have used a cut factor of 0.30 to compute the curves in Figure 2.



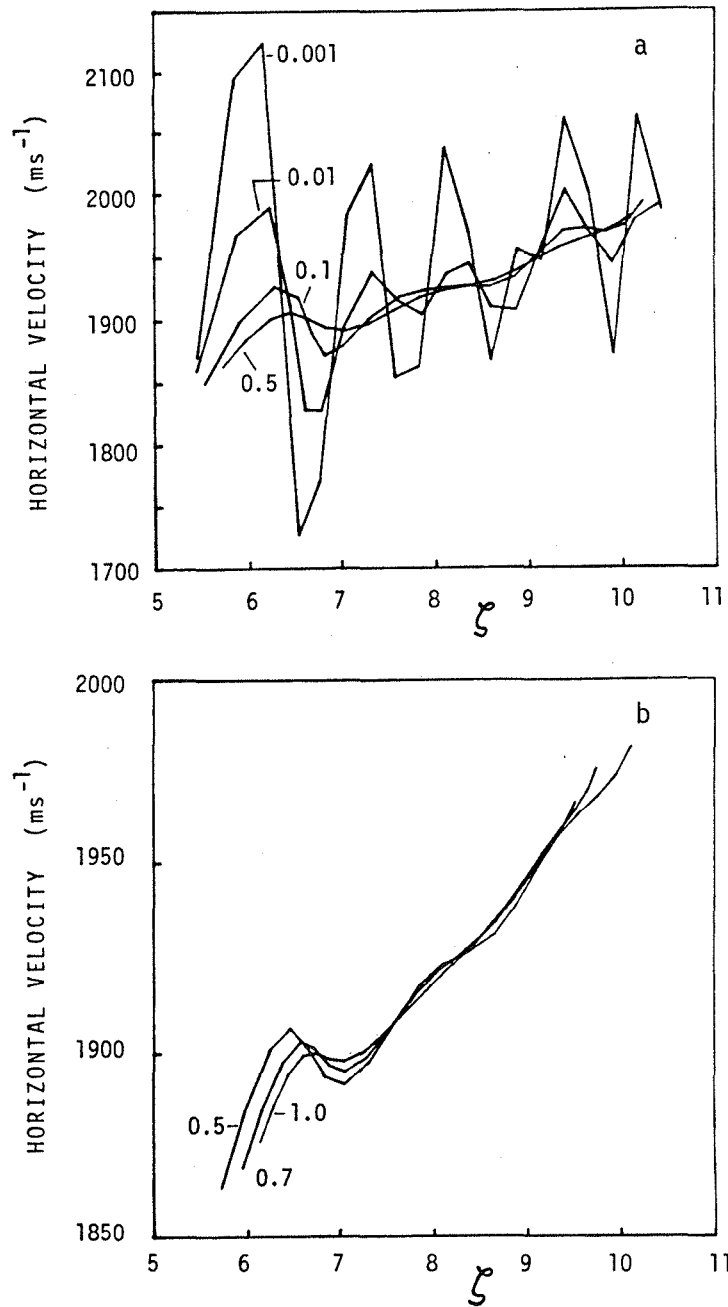
difference in pixels between the final ridge fit and the final raw medians. This error varies between 0.1 and 0.6 pixels, being higher for the higher-order and lower-degree modes. To this is added an error that is linear in frequency and is taken from our estimates of residual scale uncertainties (Hill, Toomre and November 1983). When these errors are translated to velocities, they range from  $5 \text{ ms}^{-1}$  to nearly  $600 \text{ ms}^{-1}$ . The noise in the inferred velocity is found by combining the raw errors using the same coefficients  $\alpha_1$  that give the velocity curve. For our standard case, with  $\theta = 0.5$ , the resulting error in the velocity ranges from 7 to  $34 \text{ ms}^{-1}$ . The error is higher at greater depths, which are sampled by the higher-order and lower-degree modes. Table 1 summarizes the range of errors in the inferred velocity for the various values of  $\theta$ .

Table 1. Errors in Inferred Velocity

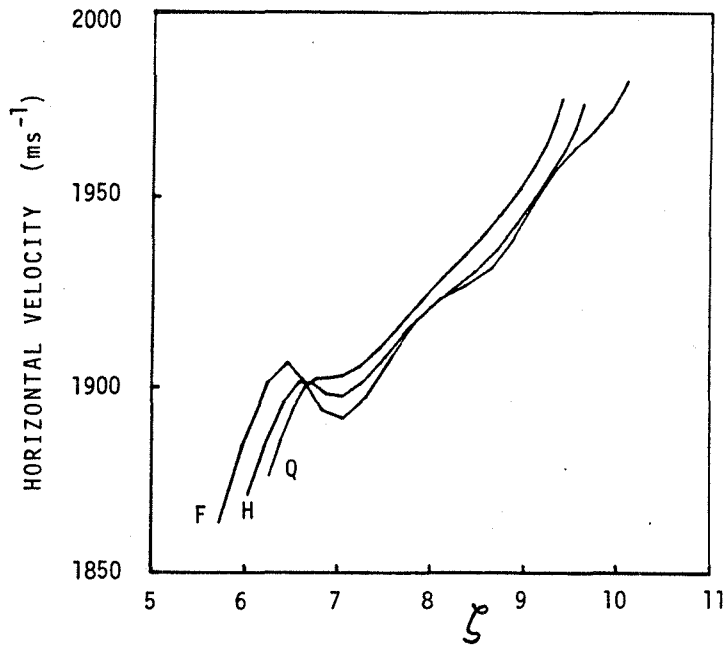
$\theta$	Error ( $\text{ms}^{-1}$ )
0	5 - 600
0.001	76 - 166
0.01	40 - 81
0.1	17 - 47
0.5	7 - 34
0.7	6 - 30
1.0	5 - 26

The choice of modes used in the inversion can also affect the resolution and accuracy of the inferred velocity. The modes, of course, must sample all the depths of interest, and increasing their number improves the resolution and decreases the final error. Figure 9 shows the effects of changing the size of the mode set. The full set is the one we have used throughout this study, and consists of 196 modes: they are approximately evenly spaced in  $\ell$  on the  $f$  and the  $p_1$  through  $p_8$  ridges, except that they avoid the frequency band of the chromospheric modes. The half mode set consists of 98 modes and was created by eliminating every other mode in the list of the full mode set. The quarter mode set totals 51 modes and was constructed from the half mode set in a similar manner. Figure 9 shows inversions for the three sets, and illustrates the degradation of resolution that results from the use of a lower number of modes. The gross features of the curve remain in all three cases. Again we see changes in the depths at which the optimal kernels are centered as fewer modes are used. This change is similar to that observed in Figure 8b for larger values of  $\theta$ . Of course, one would like to include all possible observed modes, but constraints of computer resources prevent it.

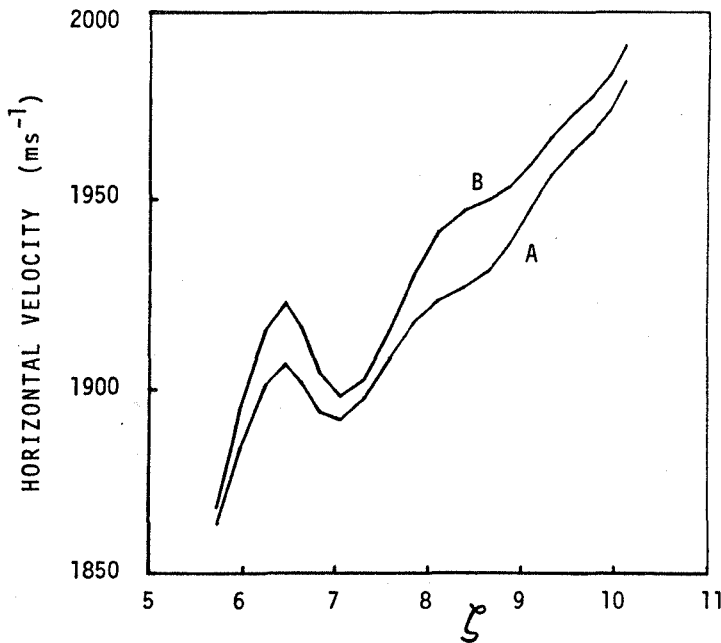
That the cancellation properties of the inversion procedure are effective in dealing with noisy data can also be demonstrated by examining the velocity curves resulting from applying the inversion to the final raw medians rather than the fitted data. This is shown in Figure 10, which compares our standard inversion obtained from the ridge fits



**Figure 8.** The effect of varying the parameter  $\theta$  in the inversion procedure. This parameter controls the amount of weight placed on the errors in the data:  $\theta = 0$  implies no weight on the errors and  $\theta = \pi/2$  implies no weight on the data. Panel (a) shows the curves for  $\theta = 0.001, 0.01, 0.1$  and  $0.5$ . A wide variation results when  $\theta$  is small. Panel (b) shows the curves for  $\theta = 0.5, 0.7$  and  $1.0$ . We have used  $\theta = 0.5$  elsewhere in our analysis.



**Figure 9.** The effect of restricting the number of modes used in the inversion. Curve F was determined using the full 196 mode set. Curve H was computed using only 98 modes, selected by rejecting every other mode from the full set. Curve Q was calculated using a set chosen by rejecting every other mode from H.



**Figure 10.** Illustration of the ability of the inversion procedure to cope with noisy data. Curve A was computed using the final fitted medians from the iterative orthogonal scheme, and curve B was calculated using the unsmoothed medians.

with that obtained from the unsmoothed medians. The curves are quite similar in shape, with that deduced from the unsmoothed medians being some 10 to 20  $\text{ms}^{-1}$  higher in magnitude. The inversion is thus quite effective in handling noisy data. We must point out, however, that these final unsmoothed medians do not have as noisy a character as the medians resulting from simply cutting the ridge in the vertical direction.

## 5. CONCLUSIONS

We have shown that our iterative ridge-fitting procedure is stable for certain choices of parameters. Moreover, the inversion procedure is quite effective in reducing errors, even when dealing with the noisy data considered here.

Of course, a reduction in the noise in the observations would be of great benefit in increasing the accuracy of the results. The increased stabilization of future instruments for measuring Doppler shifts will be of considerable advantage. Another obvious gain would come if one were to observe from space, where the data would be free from atmospheric seeing. A certain degree of improvement could also be obtained from the longer intervals of continuous observation that would be obtained, either from a spacecraft in a continuously sunlit orbit or from a ground-based network, but that improvement might only be moderate because the observing interval would still be limited by the evolution timescale of the giant cells.

## ACKNOWLEDGMENTS

The National Solar Observatory is operated by the Association of Universities for Research in Astronomy, Inc., under contract AST 78-17292 with the National Science Foundation. The Joint Institute for Laboratory Astrophysics is operated jointly by the University of Colorado and the National Bureau of Standards. The research was supported in part by the Air Force Geophysical Laboratory through contract F19628-83-K-0008 and by the National Aeronautics and Space Administration through grants NSG-7511 and NAGW-91 to the University of Colorado. Substantial support was also provided by the U.K. Science and Engineering Research Council.

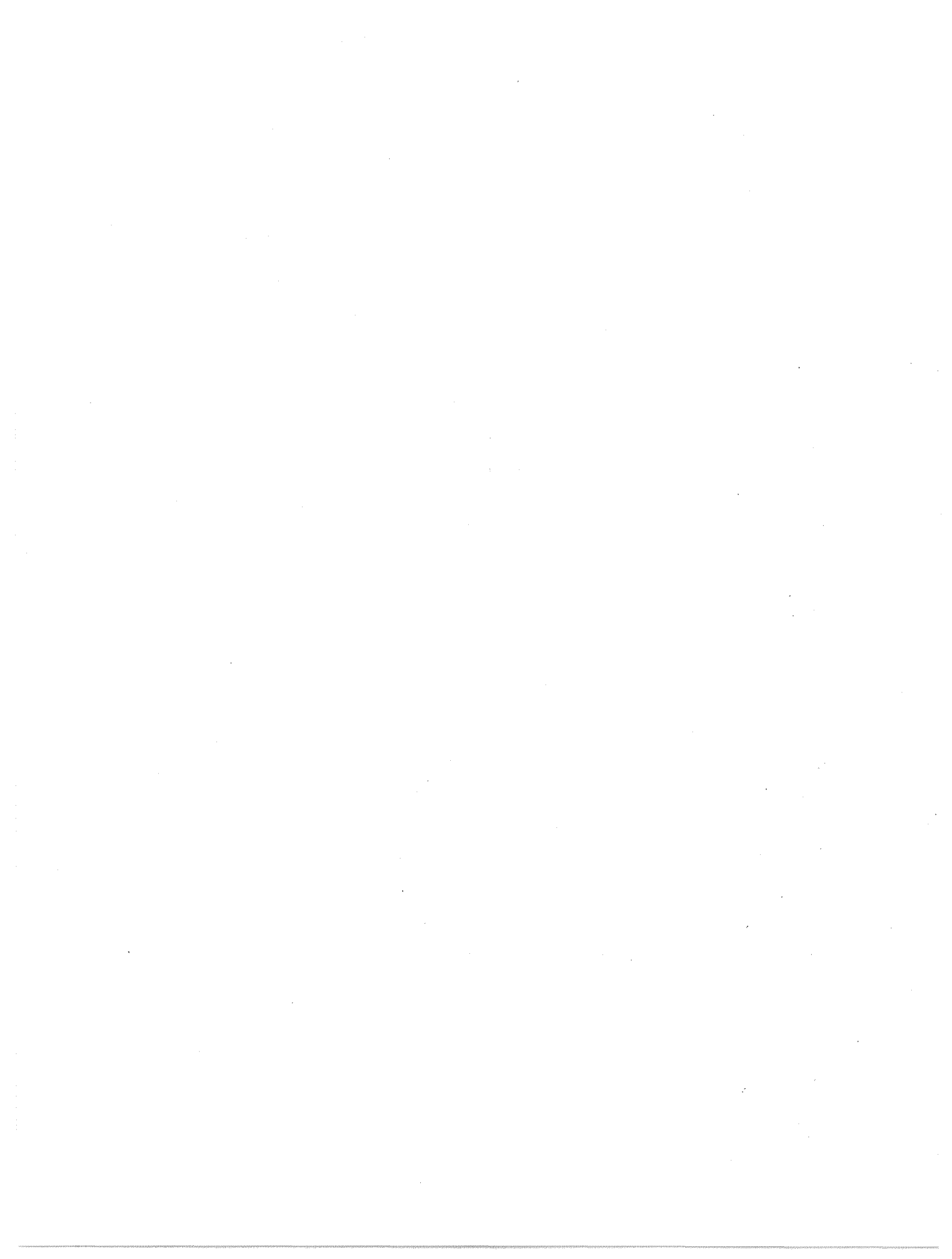
## REFERENCES

- Backus, G., and Gilbert, F. 1968, *Geophys. J. R. Astron. Soc.* **16**, 169.
- Backus, G., and Gilbert, F. 1970, *Phil. Trans. Roy. Soc.* **A266**, 123.
- Christensen-Dalsgaard, J., and Gough, D.O. 1984, these Proceedings.
- Deubner, F.-L., Ulrich, R.K., and Rhodes, E.J. Jr. 1979, *Astron. Ap.* **72**, 177.
- Duvall, T.L. Jr. 1982, *Nature* **300**, 242.
- Gough, D.O. 1978, *Proc. Workshop on Solar Rotation* (ed. G. Belvedere and L. Paterno, Univ. Catania Press), p. 255.
- Gough, D.O. 1984, these Proceedings.
- Gough, D.O., and Toomre, J. 1983, *Solar Phys.* **82**, 401.
- Hill, F. 1984, these Proceedings.
- Hill, F., Gough, D., and Toomre, J. 1984, *Oscillations as a Probe of the Sun's Interior*, (ed. G. Belvedere and L. Paterno), *Mem. Soc. Astron. Italiana*, in press.
- Hill, F., Toomre, J., and November, L.J. 1982, *Proc. Pulsations in Classical and Cataclysmic Variable Stars* (ed. J.P. Cox and C.J. Hansen, Joint Inst. for Laboratory Astrophys.), p. 139.
- Hill, F., Toomre, J., and November, L.J. 1983, *Solar Phys.* **82**, 411.
- LaBonte, B.J., Howard, R., and Gilman, P.A. 1981, *Ap. J.* **250**, 796.
- Latour, J., Toomre, J., and Zahn, J.-P. 1983, *Solar Phys.* **82**, 387.
- November, L.J., Toomre, J., Gebbie, K.B., and Simon, G.W. 1979, *Ap. J.* **227**, 600.
- Rhodes, E.J. Jr., Deubner, F.-L., and Ulrich, R.K. 1979, *Ap. J.* **227**, 629.
- Rhodes, E.J. Jr., Harvey, J.W., and Duvall, T.L. Jr. 1983, *Solar Phys.* **82**, 111.



PART 2

INSTRUMENTS AND OBSERVATIONS





## THE MAGNETO-OPTICAL FILTER, WORKING PRINCIPLES AND RECENT PROGRESS

Cacciani, A.<sup>1</sup>, Rhodes, E. J. Jr.<sup>2</sup>

<sup>1</sup> Dipartimento di Fisica dell-Universita "La Sapienza" Roma  
<sup>2</sup> University of Southern California and Jet Propulsion Laboratory,  
California Institute of Technology

### - ABSTRACT

Solar oscillation observations are in progress at the 60-foot Solar Tower of the Mt. Wilson Observatory. We present our results in a separate communication of this meeting. This note is devoted to a brief additional description of the working principle of the Magneto-Optical Filter (MOF) and a few possible optical arrangements.

### - WORKING PRINCIPLES

The working principle of the MOF has been extensively described elsewhere (see e.g. A. Cacciani and M. Fofi, 1978). Here we shall describe it in a simpler way, avoiding mathematical formulae and using instead simple physical considerations.

We consider first an incoming plane wave of non-monochromatic, non-polarized radiation which is incident on a gas or vapour cloud (e.g. Na) which is in turn strongly absorbing at a certain resonance wavelength,  $\lambda_r$ . If the optical depth is high enough, there will be no radiation transmitted at  $\lambda_r$  and the cloud will be completely opaque. (For the purpose of building up an imaging instrument, the monochromatic radiation which is isotropically scattered shortly after the absorption process is of no utility, and so we will ignore it for the moment.)

Let us now suppose that at the given monochromatic wavelength,  $\lambda_r$ , the cloud can absorb only one kind of polarization state, as is the case when the vapour is placed in a magnetic field (Zeeman effect); then the cloud will become transparent at  $\lambda_r$  since the conjugate state of polarization cannot be absorbed. (We also note that an electric field, due to the Stark effect, can produce a similar behaviour.)

Only if the incoming radiation is polarized in the same state as the absorption, will there be complete extinction. All of the other cases correspond to a partial transmission (maximum 50%), with the transmitted intensity being due to the light which is polarized in the conjugate state from that of the absorption.

Next, we note that the behaviour of a normal polarizer is similar for non-monochromatic light. In this case, an absorbing gas cloud in a magnetic or electric field can be thought of as a polarizer for the Zeeman- or Stark-split wavelength  $\lambda_r$ .

Of course the stabilization of the chosen position is also important, but this need not be very stringent because the Doppler or magnetic signal is obtained performing a pixel by pixel difference between the two wing signals divided by their sum. Small symmetrical shifts will not introduce systematic velocity drifts into the measurements.

During the spring of 1981 and the summer of 1982 one of us (A.C.) spent a few months at the Nice Observatory (by a joint Italian CNR and French CNRS agreement) testing and evaluating the filter performance with E. Fossat's group. A first  $k$ - $\omega$  diagram was obtained (J. Blamont *et al.* 1982) and a quantitative comparison of the MOF and the scattering technique was performed having in mind a possible stellar experiment (E. Fossat *et al.* 1982). In particular the stability test gave very good results. During 5 hours of a very clear day, in September when the solar line is centered within the filter wavelength, the ratio of transmitted-intensity to input-intensity varied by only 4% without any particular temperature stabilization. This corresponds to about a 10mÅ symmetrical drift in the two transparency window positions. Further measurements performed at Mt. Wilson gave a shift of a few mÅ per watt of the heater power. This is a very small effect which is able to change the amount of the transmitted light but which is almost completely ineffective in changing the Doppler signal.

#### - POSSIBLE MEASUREMENTS WITH A SIMPLE (SINGLE-CELL) FILTER

A simple MOF is composed by a glass cell containing a suitable chemical element, a little magnet of  $\sim 1$  KG, two calcite polarizers, a prefilter (30 Å pass-band) and additional optics to fit the telescope characteristics. It transmits simultaneously the two wings of the solar line. In the next section we will describe how to separate them; in this section let us briefly review what measurements are possible with the simple filter.

When operating at a low optical depth this filter selects the central part of the solar line. If the glass cell is filled with Sodium the filtered image shows chromospheric structures: faculae, supergranules, flares, Ellerman bombs and velocity-originated mottling (Fig. 2). A comparison during August, 1983, of Na images produced by the MOF and by the Universal Bi-refringent Filter (UBF) of the Big Bear Solar Observatory demonstrated the superiority of the MOF due to its narrower bandpass (50-80 mÅ instead of the 250 mÅ of the UBF), which allows it to transmit just the chromospheric core of the Na D lines.

The single cell filter is also sensitive to velocity fields in the sun. This sensitivity was demonstrated at the Nice Observatory by scanning the solar disk in the E-W direction on an entrance aperture of  $\sim 200$  arcsec.

In the case of a longitudinal magnetic field, the Zeeman effect states that the gas cloud will behave as a circular polarizer, being lefthanded for the  $\sigma^+$  components and righthanded for the  $\sigma^-$  components ( $\pi$  components are absent in this case). In the case of a transverse magnetic field it behaves as a linear polarizer, parallel to the magnetic field for the  $\pi$  components and perpendicular for the  $\sigma$  components.

In addition to the above well-known effect in the absorption spectra of Zeeman-split lines, another effect operates while the light travels across an absorbing cloud in a magnetic field. This is less widely known because it reveals itself only when the input radiation is polarized, changing its polarization in a complicated way in a narrow band in and around the absorption wavelength,  $\lambda_r$ .

This second effect does not introduce additional absorption but rather is related to the behavior of the refraction index which accompanies the absorption coefficient. In the same way that the absorption coefficient is split by a magnetic field so also is the refraction index split by the magnetic field. The result is a retardation effect between any two split conjugate polarizations.

For a detailed analytical formulation of the above two effects (magneto-optical effects) see the above-mentioned paper and Agnelli, et al. (1975) with references therein. Here it is sufficient to state that the absorbing gas cloud in a magnetic field acts as both:

- a) - a narrow-band polarizer; and
- b) - a narrow-band retardation plate.

Having this in mind it is then easy to build up a narrow-band filter. In fact, no light is transmitted by two crossed polarizers unless a third polarizer or a retardation plate is located between them. In the MOF the third element is a vapour in a magnetic field.

This brings us to the difference between the MOF and the scattering devices used by the Nice and Birmingham groups. The scattering devices take advantage of the change in the propagation state, while the MOF takes advantage of the change in the polarization state; the scattered light is spread over the entire  $4\pi$  solid angle and only a minor fraction of it can be collected. The MOF, on the other hand, gathers all of the light in the forward direction. In addition, the second magneto-optical effect (i.e. the retardation effect) is highly efficient, substantially increasing the amount of transmitted light.

#### - SELECTING WAVELENGTHS AND TUNING

The filter as described thus far accepts in principle any of the Zeeman sensitive resonance lines. Experiments are now in progress in order to accept Sodium and Potassium simultaneously.

Individual lines can, however, be selected by using interference prefilters. Normally we use a flat-top, 30Å-passband prefilter which is relatively insensitive to thermal transmission drifts. In the case of the Sodium D lines, only 6Å apart, it may be convenient to employ an alternative technique in order to avoid having to thermally stabilize such a narrow-band filter. Instead we can also use a calcite retardation plate of suitable thickness ( $\sim 1.7\text{mm}$ ) located after the exit polarizer but at a  $45^\circ$  angle with it. This plate creates an angle of  $90^\circ$  between the polarization directions of the two D lines so that they can more easily be discriminated.

The MOF is a narrow-band filter with its central wavelength perfectly stable. This is its major advantage but it is also a limitation if tuning across the solar line profile is needed.

Tuning could be achieved with the help of a variable magnetic field but our aim is to have a compact and light weight instrument suitable for space applications; hence variable electromagnets have been avoided. Actually we have adopted a permanent magnet with a fairly low magnetic field of about 1 KG, which is just the minimum field strength required for a complete separation of the  $\sigma$  components. With this magnet we can move the two narrow transmission profiles apart, into the solar line wings, by increasing the optical depth in the filter. This is an unexpected and very important result because it simulates the behaviour of the  $\sigma$  components in a variable magnetic field and allows us to control the filter's tuning by simply varying the amount of electric power provided to the sodium evaporators.

Roughly speaking what happens is the following: an increase in the optical depth also increases the absorption line width and the retardation effect in the far wings, so that in the central wavelengths the  $\sigma$  components overlap giving a dark absorption region, while in the far wings the filter becomes more and more transparent. The result is a migration towards the solar line wings of the two narrow transparency windows (Fig. 1). Note further that these windows are not the  $\sigma$  components.

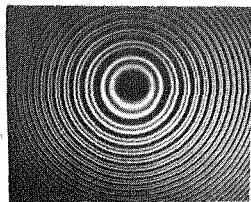


Figure 1. Fabry-Perot transparency spectrum of the MOF with 1 KG magnetic field and a constant power input for the Na evaporator. The spectral range is about  $200\text{ m}\text{\AA}$  and each transmitted wavelength has about  $50\text{ m}\text{\AA}$  width.

The wavelength positions of the two transparency windows are determined by the optical depth in the filter and always remain symmetrical with respect to the central resonance wavelength.

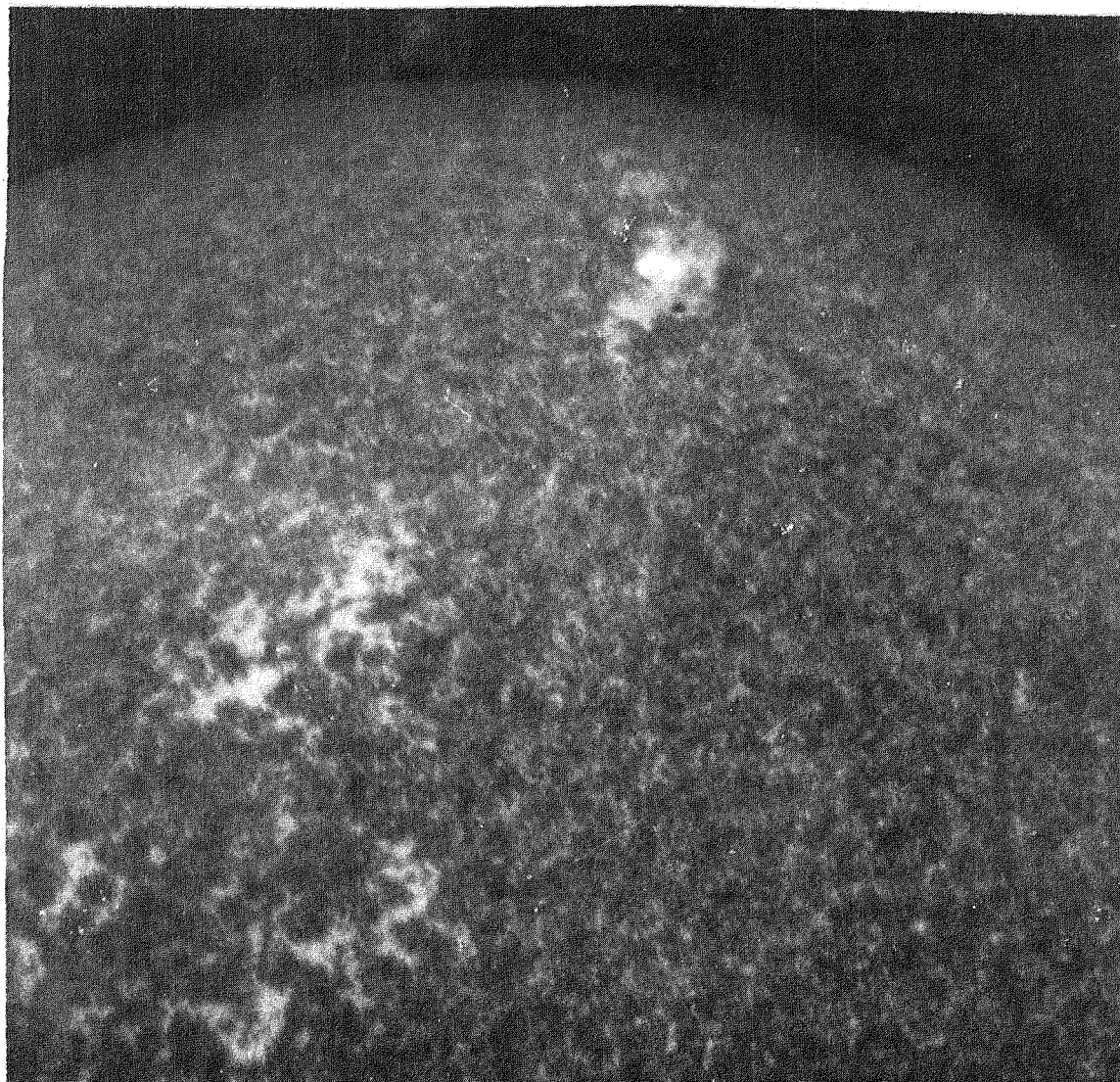


Figure 2. Na filtergram taken with the simple MOF showing chromospheric structures.

The velocity signal is position dependent on the solar disk and time dependent during the year; nevertheless it is fairly detectable (Fig. 3) and a good K- $\omega$  diagram has been obtained this summer at Mt. Wilson and shown in another communication at this meeting. Higher quality K- $\omega$  diagrams have been obtained using a two-cell system (Rhodes, et al., 1984, these proceedings).

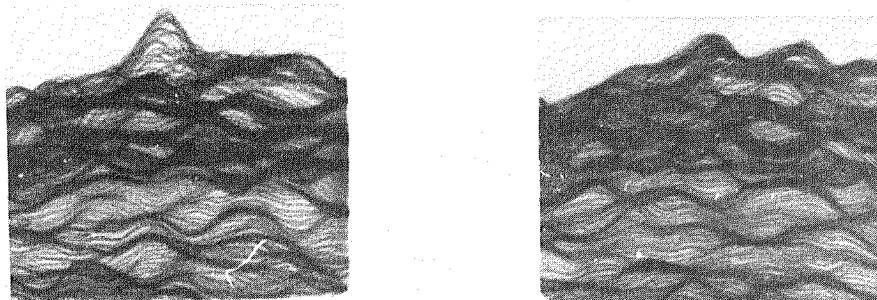


Figure 3. Two portions of a filtergram movie (successive frames being 2.5 min. apart) displaying the velocity signal in the y direction.

The simple filter can also be used to compare the velocity signal coming from two different solar regions (e.g. the Northern and Southern hemispheres) versus time (Fig. 4) (Cacciani, et al. 1981) and to observe flares in Sodium D lines (Cacciani, et al. 1980).

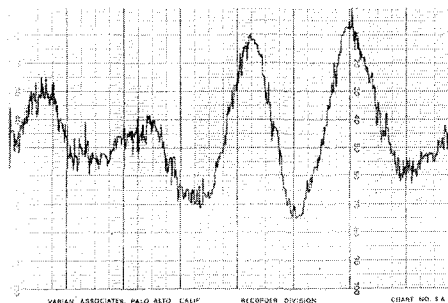


Figure 4. Record of five-min. oscillations taken with the apparatus described in Cacciani, et al. (1981). The large amplitude waves correspond to  $\sim 15$  m/s peak-peak.

## - THE TWO-CELL SYSTEM

The best way however to detect velocity and magnetic fields is to get a separation between the two monochromatic images obtained with the narrow spectral windows of Fig. 1.

There are several different ways to attain this result. All of them require an additional element called the wing-resolving element. Two different methods are currently under evaluation; the observations we are carrying on at Mt. Wilson make use of a second Na cell in a longitudinal magnetic field, located in front of the filter and without additional polarizers.

As previously stated this second cell works as a circular polarizer, lefthanded for  $\sigma^+$  and righthanded for  $\sigma^-$ , so that by using a  $\lambda/4$  plate it is possible to discriminate between the two line wings (Fig. 5).

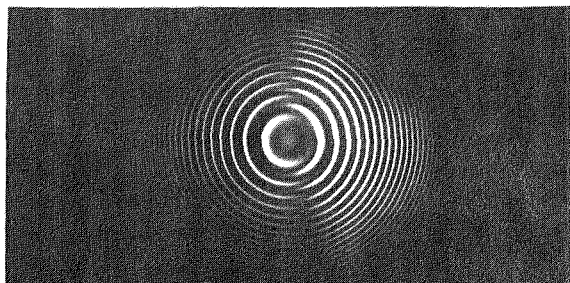


Figure 5. Fabry-Perot Interferogram showing one spectral wing in each half of the picture.

We note that the first filter polarizer of the two-cell system can be replaced by a polarizing beam-splitter yielding two simultaneous beams each filtering one wing of the solar line profile (Cacciani 1981). This arrangement however is useful only when image resolution is not required as is the case of star or sun-as-a-star pulsations, since it is very difficult to avoid differences and geometrical distortions between the two images and the two bidimensional detectors.

In any case it is possible to use one channel sequentially for spatially-resolved observations and to devote the other channel to integrated solar light measurements using a more efficient technique for recovering the signal from the noise, such as a lock-in amplifier.

In addition, the two-cell system can give magnetic maps of the sun without interfering with the velocity measurements. In fact the wing-resolving cell is oppositely polarized in the blue- and red-wings of the line profile just as the incoming radiation from a solar plage is. Hence there will be complete absorption or complete transmission of the solar magnetic polarization depending on the direction of the magnetic

field. The signal is twice that normally given by a conventional magnetograph and can be modulated by placing a KDP cell in front of the whole system. This magnetograph mode of operation was recently tested at the Big Bear Solar Observatory with the assistance of Dr. H. Zirin and his staff.

- SUMMARY

A device has been described which allows simultaneous magnetic and velocity measurements (in both imaging and non-imaging modes) without the need for a spectrograph. In this way the stability and alignment problems of the spectrograph are completely overcome. Its major advantages are: wavelength absolute reference and stability, high signal-to-noise ratio and independence of the transmission profile from the incidence angle of the solar beam. It is an imaging instrument allowing high-wavenumber analysis in the solar oscillation spectrum and a continuous monitoring of the image position through the chromospheric facular structures.

The apparatus we have been using at Mt. Wilson is assembled in a modular form. The most important part of it is a glass cell containing the Sodium vapour. The filter is easy to use but the cell is not easy to construct in an optimal way. The technology is in progress both to use Na and K together and to prevent the windows from becoming coated during a long-term operation.

- REFERENCES

- Agnelli G., Cacciani A., Fofi M. (1975) "The MOF I: Preliminary observations in Na D lines" Solar Physics 44, 509.
- Blamont J., Cacciani A., Rhodes E. Jr., Ulrich R.K. (1982) "Observation of the sun and of interplanetary space by 3-axis stabilized spacecraft placed at 1 A.U. from the sun at large distance from the earth" Response to ESA call - Nov. 15, 1982.
- Cacciani A., Fofi M., (1978) "The MOF II: Velocity field measurements" Solar Physics 59, 179.
- Cacciani A., Fortini T., Torelli M. (1980) "Na-light flare observations: McMath 13043 - July 1974" Solar Physics 67, 311.
- Cacciani A., Croce V., Fortini T., Torelli M. (1981) "Searching for 1=1 modes of Solar oscillation" Solar Physics 74, 543.



Cacciani A. (1981) "Magnetograph based on magneto-optical effects"  
Sp. Sci. Review 29, 403.

Fossat E., Decanini Y., Grec G. (1982) "A spectrophotometer devoted  
to stellar seismology", Instrumentation for Astronomy with Large  
Optical Telescopes, 169-174 C.M. Humphries (ed.) Reidel P.C.

Rhodes, E.J., Jr., Cacciani, A., Blamont, J., Tomczyk, T., Ulrich, R.,  
and Howard, R. (1984) "Evaluation of a magneto-optical filter and  
a Fabry-Perot interferometer for measurements of solar velocity  
fields", these proceedings.



EVALUATION OF A MAGNETO-OPTICAL FILTER AND A FABRY-PEROT INTERFEROMETER  
FOR THE MEASUREMENT OF SOLAR VELOCITY FIELDS FROM SPACE

Edward J. Rhodes, Jr.<sup>1</sup>, Alessandro Cacciani<sup>2</sup>, Jacques Blamont<sup>3</sup>,  
Steven Tomczyk<sup>4</sup>, Roger K. Ulrich<sup>4</sup>, and Robert F. Howard<sup>5</sup>

<sup>1</sup> Dept. of Astronomy, University of Southern California, and  
Jet Propulsion Laboratory, California Institute of Technology

<sup>2</sup> Dipartimento di Fisica dell'Universita "La Sapienza" Rome  
Italy

<sup>3</sup> Service du Aeronomie de CNRS, Verrieres-le-Buisson, France  
and Jet Propulsion Laboratory

<sup>4</sup> Dept. of Astronomy, University of California at Los Angeles

<sup>5</sup> Mount Wilson and Las Campanas Observatories

Abstract

As part of the study being conducted by the Solar Dynamics Observatory Science Working Group a program has been developed to evaluate the performance of three different devices as possible space-borne solar velocity field imagers. Two of these three devices, a magneto-optical filter of the type described elsewhere in these proceedings (Cacciani and Rhodes) and a molecular adherence Fabry-Perot interferometer have been installed in a newly-constructed observing system located at the 60-foot tower telescope at the Mt. Wilson Observatory. During July and August of 1983 time series of solar filtergrams and Dopplergrams lasting up to 10 hours per day were obtained with the filter while shorter runs were obtained with the Fabry-Perot. Two-dimensional  $k_{\perp}$ - $\omega$  power spectra which show clearly the well-known p-mode ridges have been computed from the time series obtained with the magneto-optical filter. These power spectra will be presented and will be compared with similar power spectra obtained recently with the 13.7-m McMath spectrograph at Kitt Peak by Duvall, Harvey, and Rhodes. Next, we will discuss our plans for continued evaluation of these instruments at Mt. Wilson and our plans for the establishment of a second observing station which will operate in conjunction with Mt. Wilson to provide nearly continuous observations with one of the instruments for a few months each year. Finally, the operation of the Fabry-Perot as a pressure-scanned narrow band interference filter will also be described.

## 1. The Need for a Compact Solar Oscillation Tachometer

A significant increase in our knowledge of solar internal structure and dynamics will occur once it becomes possible to obtain uninterrupted observations of solar oscillations for periods ranging from a few days to a few years. As the title of this conference suggests, one possible way of obtaining such continuous observational sequences would be the flight of solar oscillation measuring instruments onboard a spacecraft which is continually viewing the sun. For such a space-flight the obvious weight, size, and power limitations all dictate that a compact solar tachometer must be developed in order to replace the large, complicated spectrographs which are currently being employed on the ground to make oscillation measurements.

A basic velocity accuracy of 1 to 10 m/s (Gilman, 1979) will be required for this instrument. The 1 m/s accuracy limit has been estimated to be necessary for the detection of global eddies and large-scale, low-frequency oscillations. When specifying this accuracy limit, it is also essential to specify the measurement area on the sun to which the limit is applied. While a single-measurement accuracy of 1 m/s has not yet been obtained with ground-based instrumentation for regions of a few arcseconds across, averages of regions 4 arcseconds by 200 arcseconds obtained by one of us at Kitt Peak (Rhodes, et al, 1981) have yielded an accuracy of 30 cm/s. Thus, by effective areal averaging it is possible to do better than 1 m/s.

In addition to the ability to measure velocities on the order of meters per second, an equally stringent requirement is the temporal stability of the proposed instrument. To measure the small frequency shifts which enable us to study the subphotospheric flow patterns at the base of the convection zone, a single, continuous set of measurements will need to last at least 5 and possibly as long as 30 days. Thus, the instrument must not just have a high sensitivity to Doppler shifts, but must also possess a high wavelength stability. The wavelength of a solar absorption line must be known to 1 part in  $3 \times 10^8$  over the entire field of view at any time for at least 5 full days and probably longer.

As a portion of the study currently being carried out for NASA by the Solar Dynamics Observatory Science Working Group a program has been developed to evaluate the suitability of three different devices for such a space-borne solar oscillation tachometer. One of these devices, the Fourier Tachometer, is currently being installed at the Sacramento Peak branch of the National Solar Observatory. It is described in more detail elsewhere in these proceedings (Brown, 1984).

The other two devices are a resonance cell magneto-optical filter and a molecular-adherence Fabry-Perot interferometer. Both of these devices are currently being used to observe solar oscillations at the Mount Wilson Observatory. In this paper we will describe how both of these instruments are being tested and we will present the first solar oscillation power spectra to be obtained with the magneto-optical

filter at Mt. Wilson. (Since the conclusion of this conference we have become aware that a Fabry-Perot interferometer is also being studied for solar oscillation measurements by J. Harvey and R. Smartt at Kitt Peak; however no details of the experimental setup being employed there were presented at this conference and we are therefore unable to comment on how the Kitt Peak setup differs from that employed at Mt. Wilson.)

## 2. The Magneto-Optical Filter

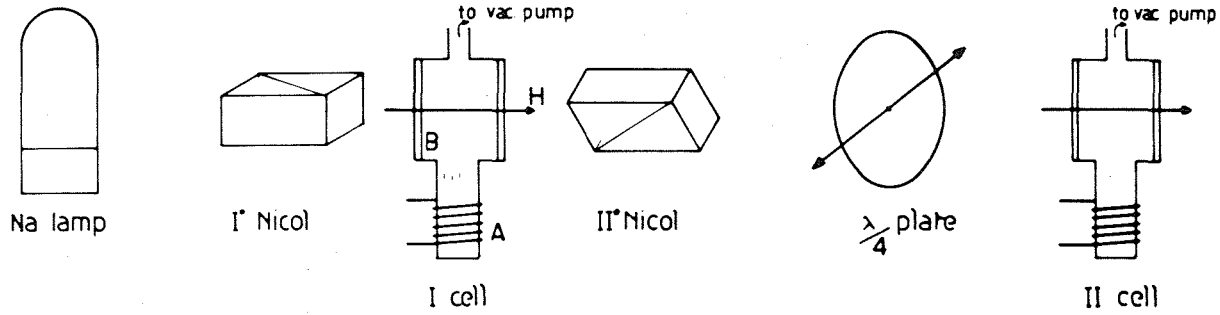
The resonance cell filter currently being tested at Mt. Wilson was developed in Italy by A. Cacciani. Its principles of operation are described more fully in the preceding paper (Cacciani and Rhodes, 1984; also see the references cited therein). Here we schematically illustrate its operation in Figure 1.

The experimental apparatus is shown at the top while the different states of polarization of the light beam are shown at the bottom. The unpolarized light from the Na emission lamp is passed first through a linear polarizer. The linearly polarized light is then passed through an evacuated chamber containing a small amount of Na vapor. Since this cell is placed in a longitudinally directed magnetic field,  $H$ , the light in two narrow bandpasses, equally spaced in wavelength on either side of the line center, is converted into  $\sigma$  left- and right-hand circularly polarized light by the Righi effect (analogous to the Zeeman effect). When the light from this cell is then passed through a second linear polarizer, this one oriented at a  $90^\circ$  angle with respect to the first, all of the linearly polarized light located outside of the two narrow bandpasses is extinguished. Additionally the light in the two spectral windows is converted into linearly polarized light. Only this light in these two bandpasses is transmitted by the second polarizer. Thus, the transmitted light from this single cell consists of two narrow spectral bandpasses. Actually, a pair of bandpasses is transmitted for both the Na D1 and D2 lines. Also, while the operation is illustrated here with an emission lamp the operation is the same for solar absorption lines. Because of the fact that light is transmitted through this first cell on both sides of each spectral line, the images obtained are sensitive mainly to temperature-induced changes in the equivalent widths of the input spectral lines.

If the linearly polarized light transmitted by the second polarizer is then passed through a quarter-wave plate the light will be converted to circularly polarized light. If this light is then passed through a second Na-filled cell embedded in a longitudinal magnetic field, only one of the two spectral windows (e.g. the one located to the "red" side of the line) is transmitted while the light in the other window (the one on the "blue" side) is absorbed by the Na vapor in the cell. By rotating the quarter-wave plate by  $90^\circ$  about the optical axis, the sense of circular polarization of the light entering the second Na cell is reversed. This reversal of polarization means that the light in the "red" bandpass which was previously transmitted is now absorbed while the light in the "blue" bandpass is now transmitted. Thus, in the two-cell mode of operation it is possible to obtain filtergrams in the light

1a

schematic  
experimental  
disposition



1b

filtering  
steps and  
polarizations

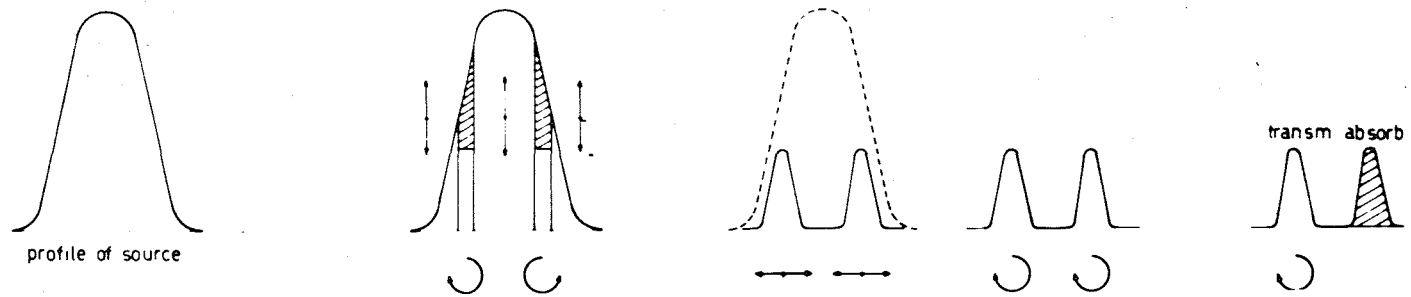


Figure 1. Schematic description of the operation of the Gacciani magneto-optical filter.

of only one of these two narrow bandpasses at a time. Furthermore, if the quarter-wave plate is continuously rocked back and forth between successive exposures, every pair of filtergrams will consist of one "redwing" filtergram and one "bluewing" filtergram. Finally, by subtracting each "red" image from its corresponding "blue" image on a pixel-by-pixel basis and then dividing by the sum of the two images in each pair, it is possible to generate a time series of Dopplergrams (i.e. velocity maps). Fourier transformation of one of these Doppler time series then results in a diagnostic ( $k_h - \omega$ ) diagram.

One of the evacuated Na cells is shown here in close-up in Figure 2. The Na is contained in the four small "pits" which project from the sides of the cell. Prior to the placement of the Na in the pits the cell is completely evacuated in order to remove atmospheric contaminants and to provide a vacuum into which the Na is evaporated when the pits are heated. Heating wires are shown on two of the four pits of this particular cell. The outer diameter of this cell is about 24 mm.

The cell is shown in Figure 3 as it is installed in a permanent magnet at Mt. Wilson. With the addition of a 30 Angstrom-wide prefilter and the two polarizers the cell-magnet combination is converted into a single-cell version of the filter. The light enters the filter from the right side of this figure, passing first through the prefilter and then through the first polarizer, the cell, and finally the second polarizer, which is hidden here at the left of the permanent magnet.

The transmission profile of such a single-cell filter is shown in Figure 4. This figure was obtained by mounting the filter in front of the entrance aperture of the spectrograph located at the Mt. Wilson 150-foot tower telescope and then mechanically scanning a photoelectric tube along the dispersion of the spectrograph. The scanning steps were each 3 mA wide so that the two transmission peaks are shown to be separated by 175 mA, while their half-widths are shown to be about 75 mA each. Somewhat to our surprise we found that the separation of the two transmission peaks can be altered by changing the amount of power applied to the heater coils of the sodium cell. In particular, the 175 mA separation shown in Figure 4 was obtained for an input current of 1.375 amps at a voltage of 10 volts, while a decrease in the current resulted in a decrease in the separation of the two peaks. Figure 5 shows that for an input current of 1.250 amperes at the same voltage the peak separation has been decreased to 100 mA. On the other hand, an increase to 1.500 amperes in the input current resulted in an increase in the separation to about 250 mA. In each case the transmission peaks moved symmetrically about the line center and retained the same widths.

From these tests we found the tunability coefficient of the filter to be on the order of one milliAngstrom of total separation per 17 milliwatts of power (or 0.5 mA per 17 milliwatts for each of the two bandpasses). Since the separation of the two bandpasses is always symmetrical about the line center, the small wavelength shifts introduced by small variations in the input power supply might not be

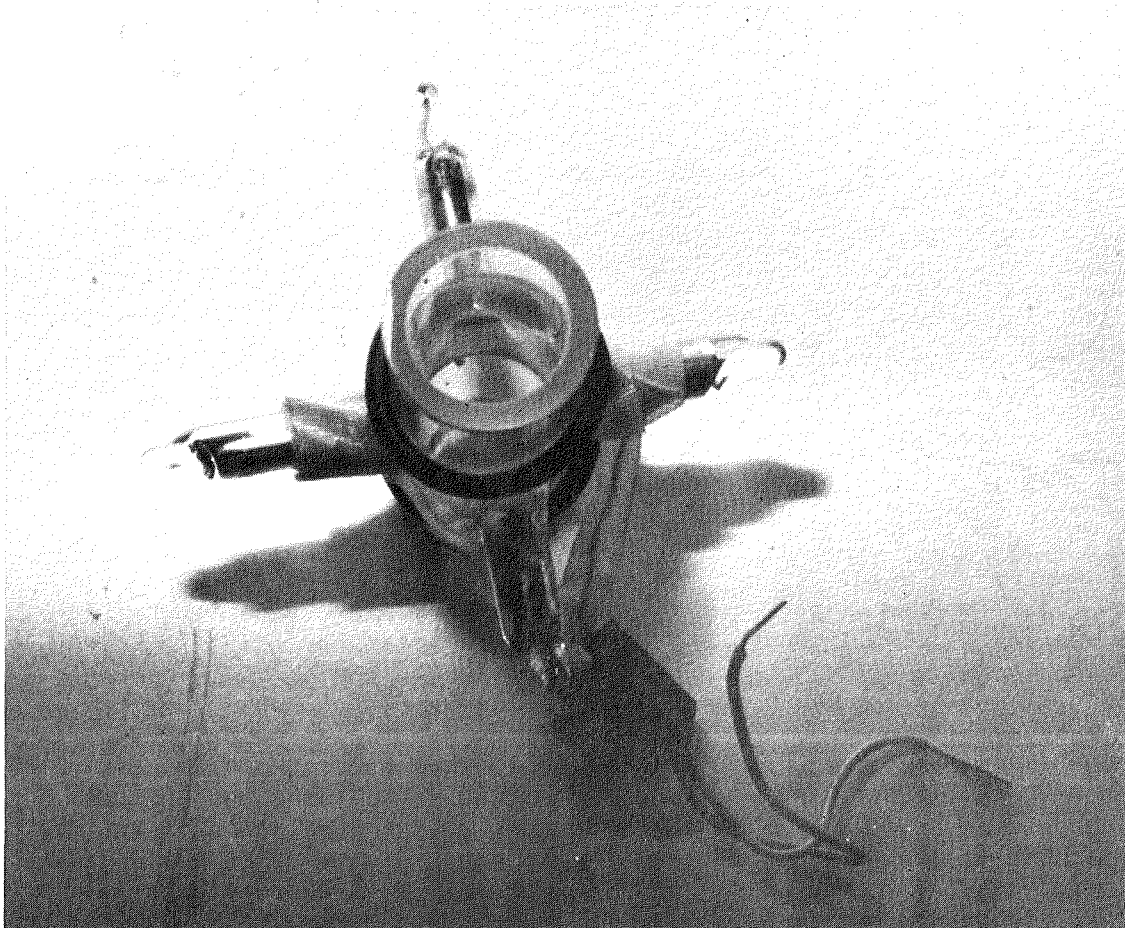


Figure 2. Close-up of a single Na-filled cell.



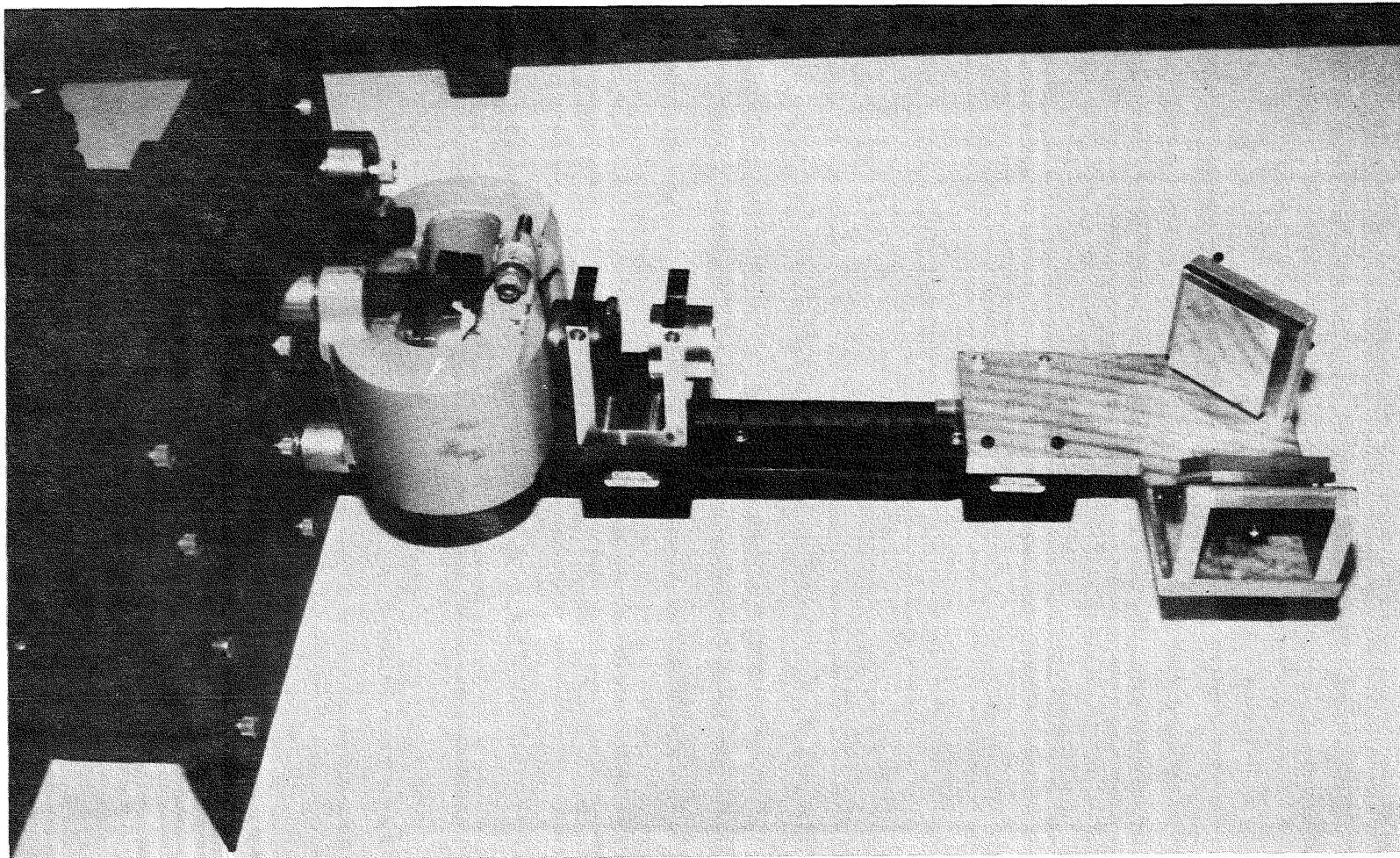


Figure 3. A single-cell filter is located at the left-center. The filter consists of a prefilter, a linear polarizer, one Na cell, and a second polarizer. The cell is located within the permanent magnet and the second polarizer is hidden to the left of the magnet. The light enters from the right.

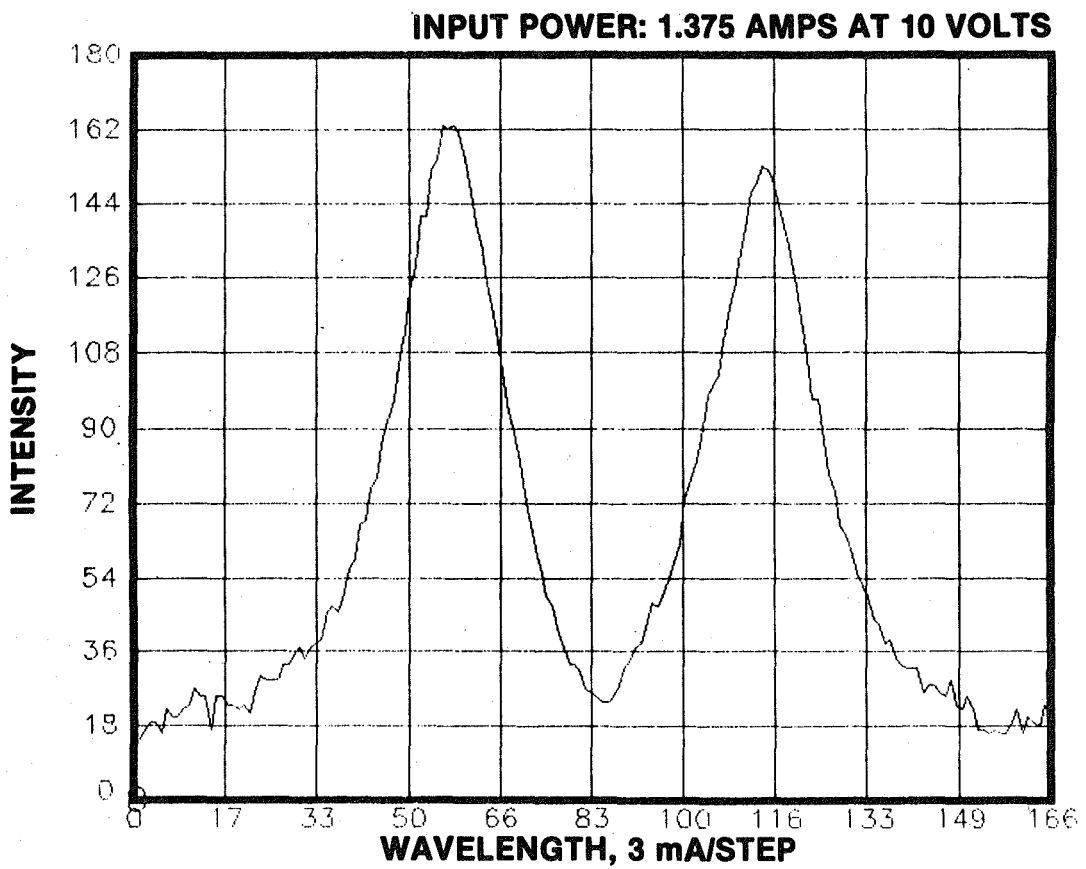


Figure 4. Transmission profile of single-cell filter. For the 13.75-watt input power shown the separation of the two bandpasses is 175 mA.

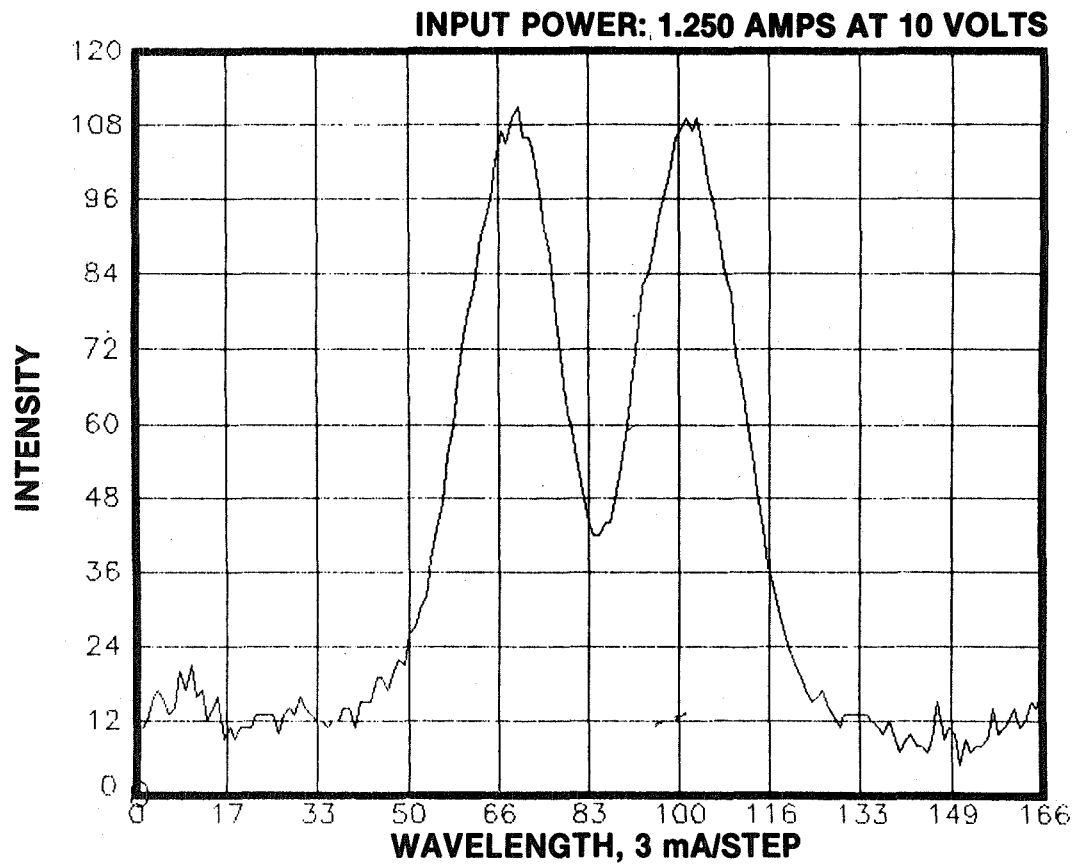


Figure 5. Transmission profile of single-cell filter for an input power of 12.50 watts. The separation of the two bandpasses has decreased to 100 mA.

expected to introduce spurious velocity drifts in the measurements. However, because the observed Na D lines are asymmetrical in the sun, such symmetrical bandpass shifts can indeed introduce small apparent velocity errors. In order to test the velocity sensitivity of the MOF to these small input power drifts we used the observed solar NA D1 line profile to obtain estimates of the actual velocity errors. We employed two 75 mA wide bandpasses each located 100 mA away from the line center. Then, for an assumed power drift of 17 mwatts (or roughly 0.14 percent of the 12.5 watt input power) we found apparent velocity errors of 0.4 m/sec for the case of zero solar Doppler shift and 0.1 to 0.9 m/sec for the case of a simultaneous 50 m/sec solar Doppler shift in addition to the power-supply-induced variations.

In fact, with current spacecraft power supply designs it is possible to obtain better than 0.1 percent power stability (even down to 0.01 percent on the short-term) and so the above numbers are upper bounds to the spurious velocity drifts to be expected in the MOF due to input power fluctuations. On the other hand, the optical depth within the MOF cells depends not just on the level of input power but also on the amount of power dissipated by the cells. Hence, it may be necessary to place the cells in a temperature-stabilized oven to stabilize the power dissipation in addition to stabilizing the input power. Tests are currently in progress to determine the sensitivity of the measured velocity to changes in the amount of power that is dissipated during MOF operation.

A filtergram of a portion of the solar disk obtained with the filter in its single-cell mode of operation is shown in Figure 6. The two-cell version of this filter is shown as it is currently installed at Mt. Wilson in Figure 7. There are two identical Na cells, one located within each of the two permanent magnets. Located between the two magnets is the moveable quarter-wave plate. The housing to the left of the second magnet contains a reversible AC stepper motor which is used to drive the Geneva movement located at the front of the motor. The Geneva movement is in turn connected to the quarter-wave plate by the long horizontal arm. By alternately driving the Geneva movement in the clockwise or counter-clockwise directions with the motor, we can move the quarter-wave plate to positions of first  $+45^\circ$  and then  $-45^\circ$  with respect to the direction of the linear polarizer. Since the stepper motor is under the control of an LSI-11, we can periodically flip the quarter-wave plate back and forth throughout each observing run in order to generate the "red" and "blue" filtergrams needed to produce Dopplergrams as was described earlier.

An example of the transmission profile for such a two-cell filter is shown in Figure 8. Here it is evident that the insertion of the quarter-wave plate has eliminated the red transmission peak that was visible in Figures 4 and 5, while leaving the shape of the blue peak relatively the same. A movement of the quarter-wave plate to the  $-45^\circ$  position would eliminate the blue bandpass while simultaneously re-transmitting light through the red bandpass.

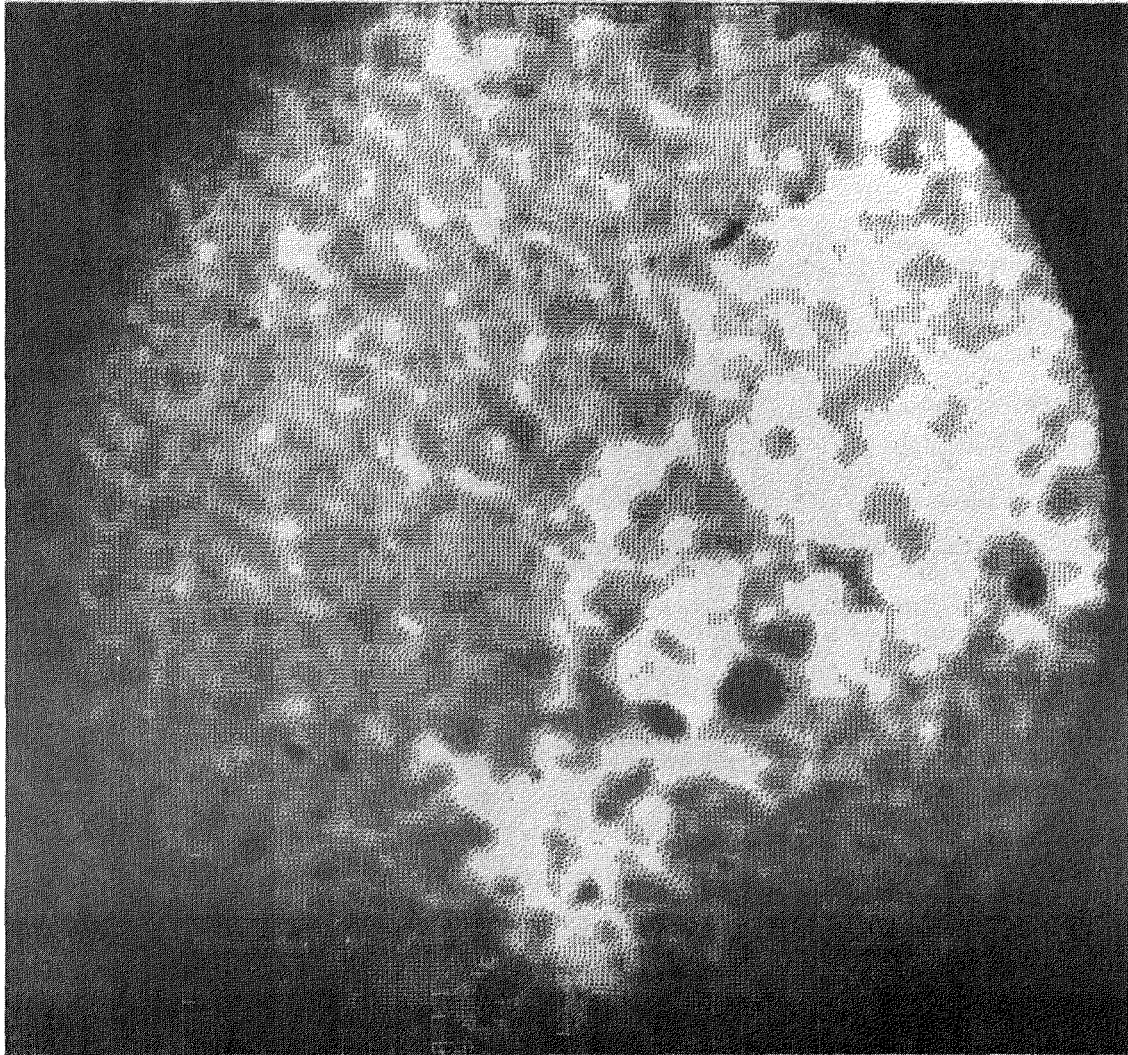


Figure 6. Filtergram of central portion of the solar disk obtained with a single-cell filter.



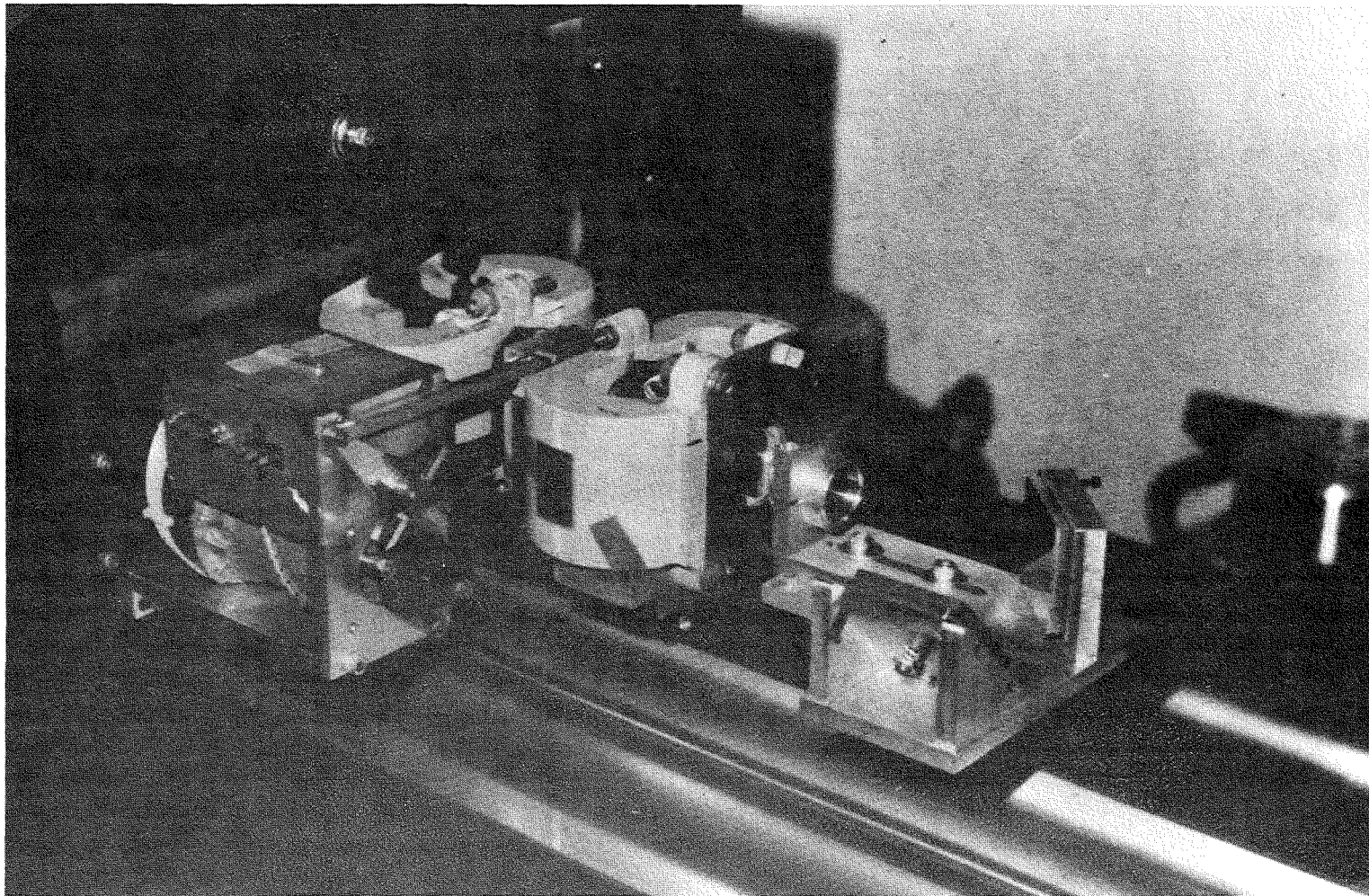


Figure 7. Two-cell version of filter as it is currently being employed at Mt. Wilson. The quarter-wave plate is located between the two magnets and can be rotated through  $90^\circ$  by the motor and Geneva movement located at the left.

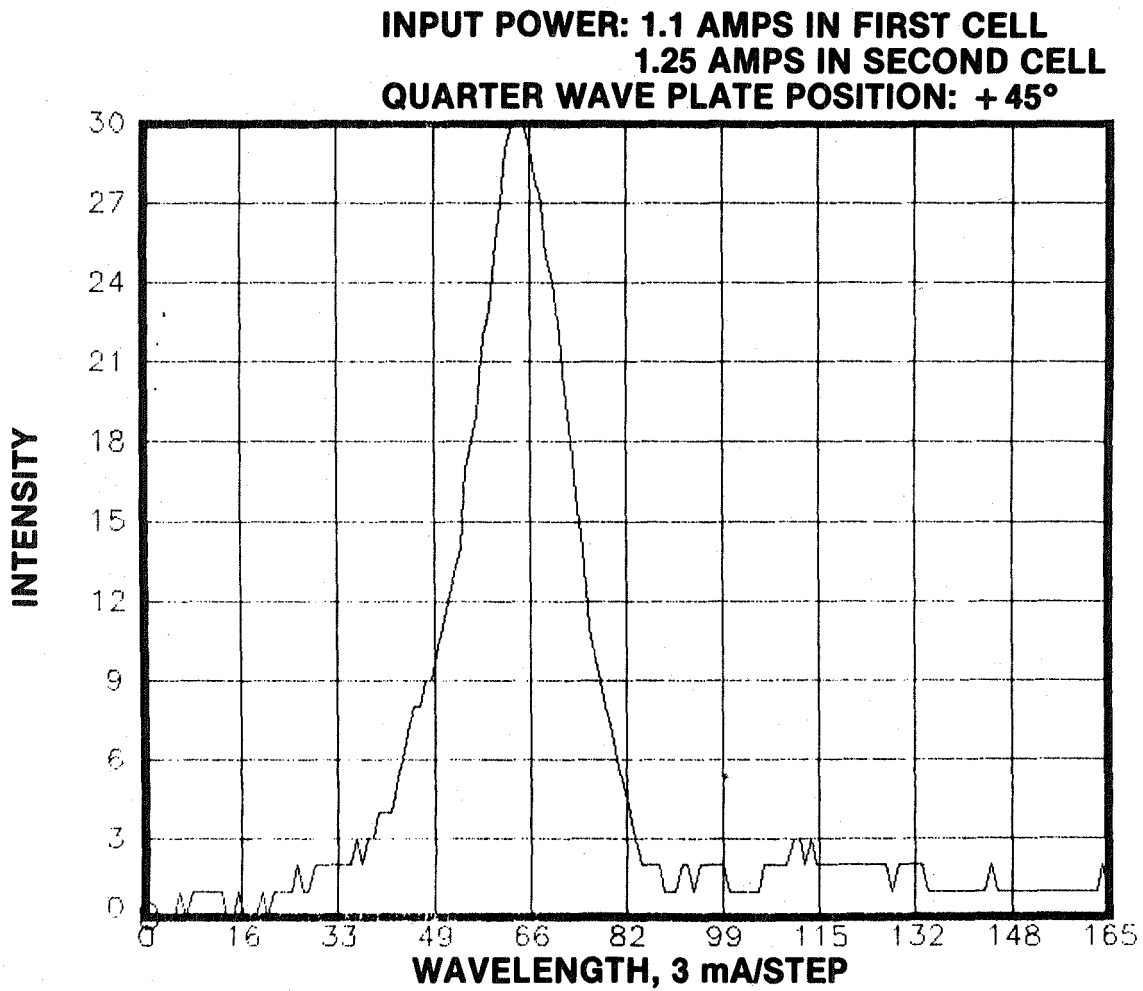


Figure 8. Transmission profile of two-cell filter. Only one of the two spectral bandpasses is transmitted at a time. Rotation of the quarter-wave plate through  $90^\circ$  results in the transfer of the bandpass to the opposite side of the spectral line.

### 3. Mt. Wilson Oscillation Observing System

The new Mt. Wilson oscillation observing system integrates CID and CCD camera's, an array processor, a 32-bit minicomputer, a large-capacity disk memory system, and custom-designed real-time telescope hardware into the spectrograph of the 60-foot solar tower telescope at Mt. Wilson. This system will be dedicated to the acquisition of solar data. The new system was recently described to IAU Colloquium No. 66, (Rhodes, et al., 1983).

First, the 60-foot tower telescope is the telescope with which magnetic fields were discovered on the sun (Hale, 1908) and with which the solar "5-minute" oscillations were discovered. The telescope is a 60-foot tall tower with a coelostat at the top. The solar image that is acquired by the coelostat is aimed vertically down the tower to either of the two existing optical systems. In the optical system used during July and August of 1983 for this project a 10-inch diameter off-axis paraboloid mirror and a swing-in secondary mirror provide a folded 30-foot focal length light path. This system forms a 3.4 inch diameter image at the entrance slit of the pit spectrograph located at the base of the tower.

The spectrograph is a vertical pit spectrograph having a 13-foot focal length. It is located within a 25-foot deep concrete pit which is located beneath the floor of the observing room.

The design of our new equipment was chosen so that the existing sequence of daily spectroheliograms which is currently being obtained at the telescope could be continued each morning and then the remainder of that day could be devoted to this oscillation project. In this design the CID camera assembly is mounted in an enclosure which is attached to the side of the main spectrograph. Through the use of a narrow periscope assembly, light can be deflected through the magneto-optical filter to the CID detector array which is in turn located within a liquid-nitrogen-cooled dewar. By allowing the whole dewar-periscope assembly to move on rails, we can position it so that it is completely out of the light which is used to expose the photographic plates placed above the spectrograph's exit slit each morning.

Other components of the system include: 1) a high-speed rotating disk shutter which provides the short (1.2 sec) exposure times needed for this project; 2) an image rotator; and 3) a guider which can scan the image in two orthogonal directions under computer control. The instrument enclosure and moveable stage described above were machined at U.C.L.A. and at the Jet Propulsion Laboratory, were assembled at the Jet Propulsion Laboratory and are now permanently mounted on the side of the spectrograph. The enclosure covers a hole which was cut into the side of the spectrograph to allow for the movement of the mechanical stage and the periscope assembly inside of it. The stage assembly was designed to provide rotational symmetry about the center-line of the downward-looking CID chip, while also providing for focus



travel and motion along the dispersion of the spectrograph. When the stage and periscope are located within the spectrograph, the light arrives at the periscope from the aperture in the disk shutter assembly which is installed in place of the normal entrance slit assembly on top of the spectrograph. The light is then reflected to a second 45° flat mirror which is located in front of the magneto-optical filter. After passing through the filter the light passes to another mirror which reflects the light upward where it is imaged upon the downward-looking CID chip in the dewar. In this fashion we can completely bypass the grating of the spectrograph during our oscillation measurements. At these times we are merely employing the spectrograph as a large light-tight enclosure for the CID camera.

The CID camera system was built by Photometrics, Ltd of Tucson, Arizona. The CID chip itself is a General Electric Model D 244 x 248 pixel charge-injection device. As previously described this chip is located in a dewar which is cooled to the temperature of liquid nitrogen to minimize thermal noise. Also part of the camera head is a small electronics box which is mounted on the dewar. This box includes the 12-bit high speed analog-to-digital converter, the analog circuits, the clock drivers, and the line drivers and receivers for the cables connecting the camera to the controller and to the array processor.

This camera is similar in overall design to the 244 x 248 pixel camera currently in operation at Kitt Peak, since it was designed by the same man, Richard Aikens of Photometrics, who designed and built the Kitt Peak system (Aikens, 1980). However, this camera has been improved over the Kitt Peak camera in several respects: 1) it is faster than the Kitt Peak system, with a pixel read time of only seven microseconds; 2) it provides in-camera row-crosstalk correction; 3) it contains enough memory to store a single bias frame (i.e. a low-intensity, artificially-exposed frame) within the camera; and 4) it subtracts the bias frame pixel-by-pixel from each data frame before outputting the digitized pixels.

The CID camera system is controlled by an LSI-11 microcomputer which functions solely as a control microprocessor. The LSI-11 obtains interrupt pulses from the rotating cylindrical shutter and commands the camera to alternately build up exposures or read them out. It also commands the bias-frame exposures and flat field exposures. The LSI-11 also controls the guider whenever it is used to move the solar image.

From the CID camera the data passes through an auxiliary I/O port into a high-speed CSPI MAP 300 array processor. The camera is designed to signal the I/O device upon the completion of each exposure and to signal it as each pixel within a single frame is ready to be output to the array processor. The CSPI MAP 300 processor is designed for high-speed asynchronous applications and is configured with memory on three different internal buses. Thus, the CID camera system can send one picture of data to the MAP 300, while the processor is computing with the previously-transmitted picture's data. This double-buffered approach allows for a very high speed, yet it is completely synchronized

to the operation of the camera on a frame-by-frame basis.

The array processor divides each input filtergram by a previously-stored flat-field exposure on a pixel-by-pixel basis. In the case of a single-cell filter observing run these corrected filtergrams are then sent to a Systems Engineering Laboratories (SEL) 32/77 computer before storage on one of three 300-megabyte disk storage subsystems. The SEL 32/77 is a 32/bit superminicomputer containing 1.25 megabytes of 600-nanosecond MOS memory. During each observing day the SEL 32/77 routes the filtergrams directly to the disks and then at the conclusion of each day's observations, the SEL reads the data back from the disks, converts the three-dimensional (x,y,time) arrays into two-dimensional (x,time; or y,time) arrays and then computes the two- or three-dimensional power spectra which result from each day's worth of data.

#### 4. July-August 1983 Magneto-Optical Filter Results

The first data obtained during July 1983 was obtained with the single-cell version of the filter. Several series of filtergrams similar to that in Figure 6 ranging in length up to 9 hours were obtained. The intensities within a 128 x 128-pixel square at the center of the illuminated circle were averaged along columns. Thus, each filtergram was converted into a set of 128 average intensities. This averaging process was then repeated for all of the filtergrams within a given run. The resulting (x,t) array of intensities was then transformed into a two-dimensional  $k_h$ - $\omega$  power spectrum. The power spectrum which resulted from one such time series is shown here in Figure 9. The solar p-mode ridges are present in this spectrum; however, as we will shortly demonstrate they are not as distinct from the background noise as they are when velocity measurements are used in place of intensity measurements.

After completing these single-cell runs we installed the second cell and the quarter-wave plate as was shown in Figure 7 and began to acquire time series of alternating red and blue filter-grams. As with the single-cell images the data was first transferred to the MAP 300 array processor for flat-field correction. Then in some cases the corrected filtergrams were stored on the disks while in other cases each pair of filtergrams was converted into a Dopplergram with the following simple formula:

$$\text{velocity} = \text{constant} * \frac{I_{\text{red}} - I_{\text{blue}}}{I_{\text{red}} + I_{\text{blue}}}$$

where  $I_{\text{red}}$  and  $I_{\text{blue}}$  refer to the intensities of a given picture element in the red and blue filtergrams, respectively.

A sample Dopplergram which resulted from this operation is shown in Figure 10. This Dopplergram covers an area on the sun of about 800 arcseconds in diameter. The rotation of the sun is visible here as the transition from positive to negative velocities across the image.

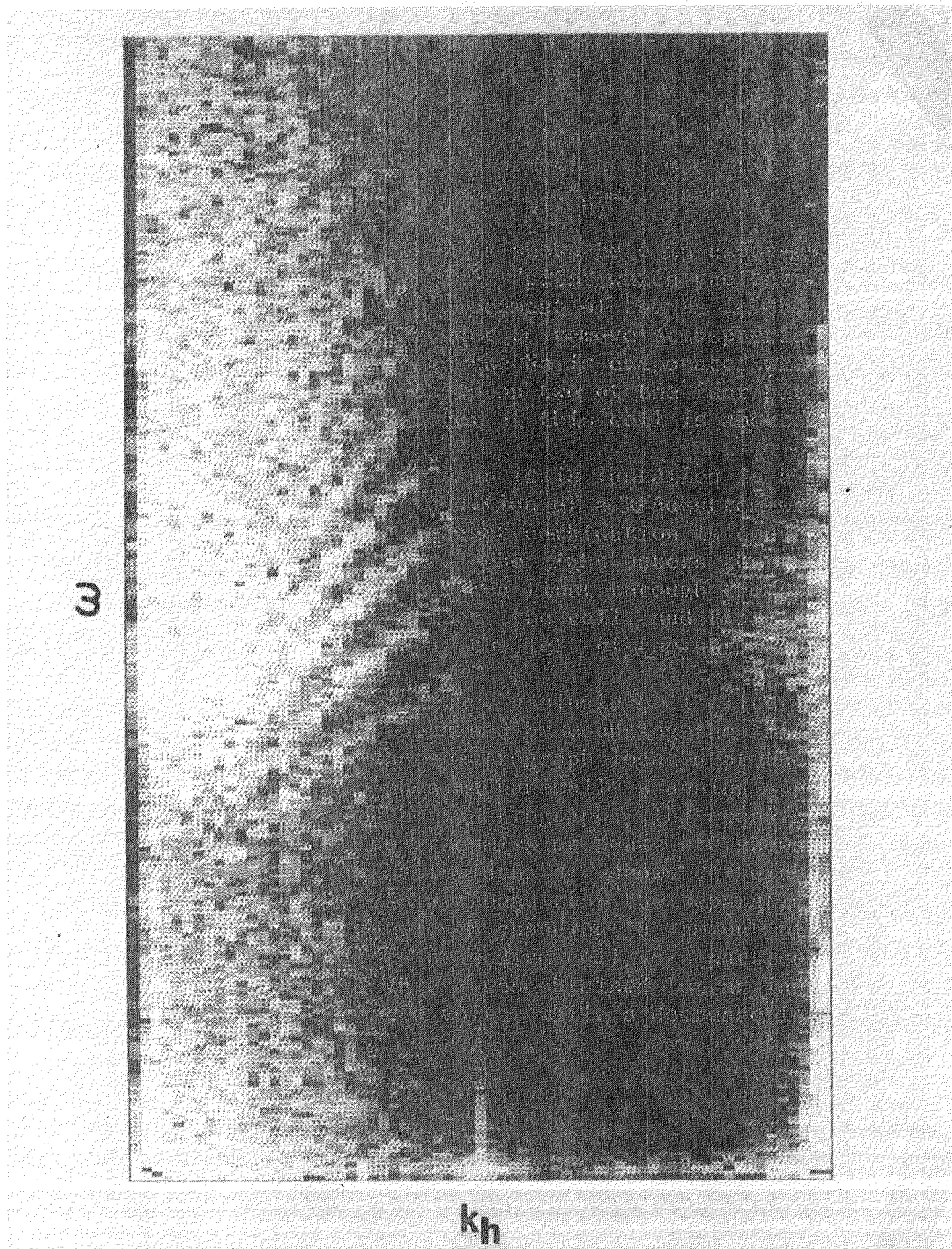


Figure 9. Two-dimensional ( $k_h$ - $\omega$ ) power spectrum obtained from a 9-hour time series of filtergrams obtained with single-cell filter operation (i.e. only intensity information was used in computing this power spectrum).

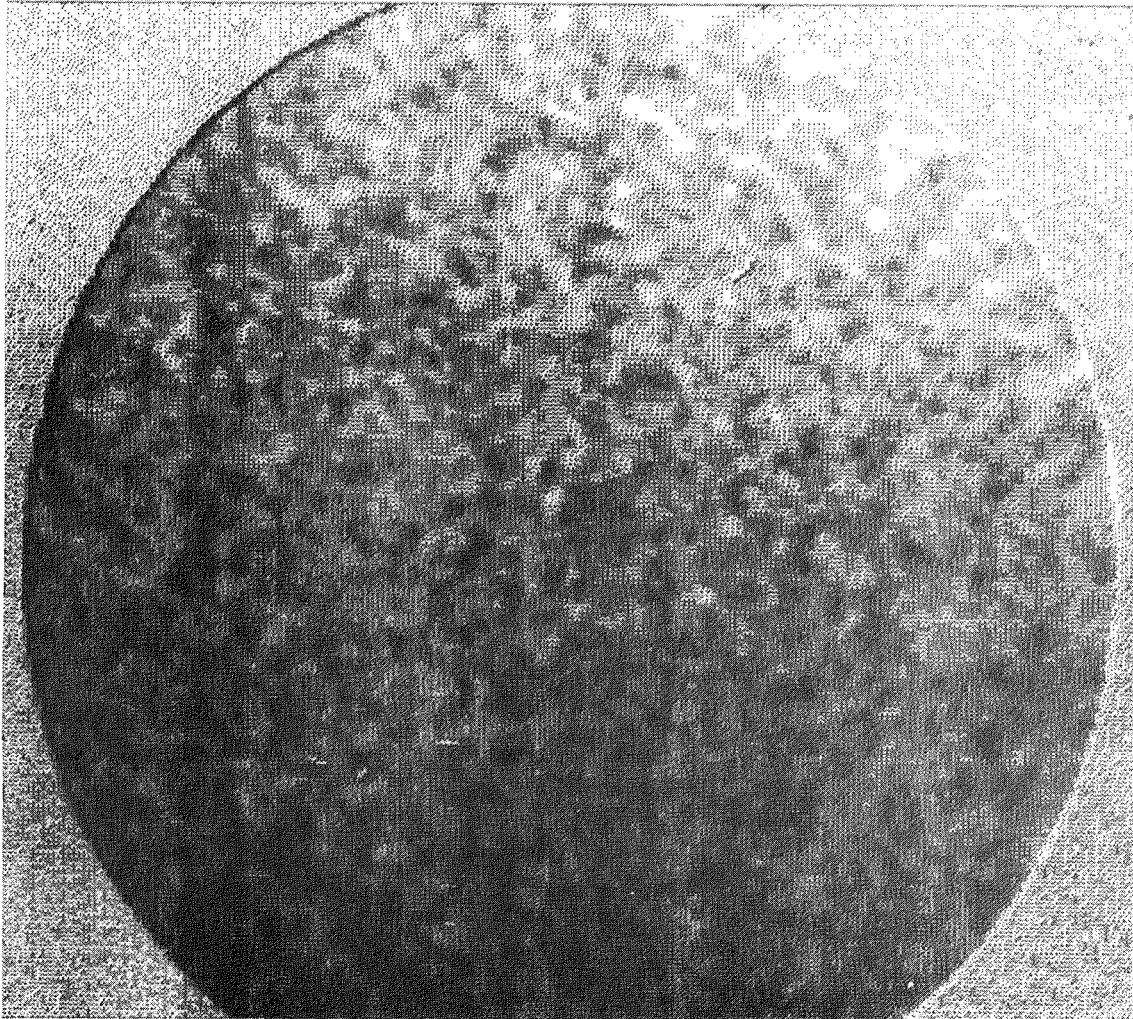


Figure 10. A Na D Dopplergram obtained by differencing a "red" and a "blue" filtergram obtained with the two-cell mode of operation. This Dopplergram covers only the central 800" of the solar disk but solar rotation is clearly visible as the variation in mean gray level running from lower left to upper right.

In order to study the linearity of the MOF over the 4 km/sec dynamic range introduced by solar rotation, a series of several Dopplergrams similar to that of Figure 10 was obtained at a series of different positions across the solar disk. The intensity differences in each Dopplergram were then averaged and then resulting average values of  $I_R - I_B / I_R + I_B$  were plotted as a function of position across the sun. One of the first plots obtained in this fashion is shown here in Figure 11, where the "E" and "W" refer to positions near the east and west limbs respectively. Overall the MOF appears to be linear over the required dynamic range (the small wiggles are mainly due to imprecise positioning of the various images across the solar disk in this first hand-positioned run).

Evidence that the MOF is in fact sensitive to actual line-of-sight velocities between Mt. Wilson and the sun is shown in Figure 12 where we plot the mean values of  $I_R - I_B / I_R + I_B$  from 1024 successive Dopplergrams. The low-degree "5-minute" oscillations are visible in the mean intensity differences and also the earth's rotation is visible as the curve during the day. A calibration program is nearing completion at Mt. Wilson which will soon enable us to compute the sun-earth radial velocity component at the time of each Dopplergram. Then we will perform a linear regression analysis of  $I_R - I_B / I_R + I_B$  upon the computed velocity in order to determine the slope of the  $\Delta I$ -vs-radial velocity curve. This slope will then enable us to convert measured intensity differences directly into radial velocities.

In order to obtain  $k_h$ - $\omega$  power spectra from the time series of Dopplergrams we obtained with the MOF we first spatially averaged the intensity differences (since we didn't yet have the calibration coefficient available) in a single Dopplergram along its columns. This spatial-averaging process resulted in 128 column-average "velocities" from each Dopplergram. A sample of the spatial temporal "velocity" maps which resulted from applying this frame-by-frame averaging process over an entire day's observing run is shown in Figure 13. The 128 column-averaged intensity differences ("velocities") from each frame are plotted using a set of gray levels. Since time runs in the direction illustrated by the arrows, it is possible to follow how the radial velocity varied at each spatial point by tracking that point in the direction given by the arrows. The "five-minute" oscillations show up in this format as the alternating white and black stripes throughout the run.

Each (x,time) "velocity" data set was then converted directly into a  $k_h$ - $\omega$  power spectrum with the same program used to convert the "intensity" dataset into Figure 9. Two such "velocity"  $k_h$ - $\omega$  power spectra are illustrated here in Figures 14 and 15 (although Figure 14 is actually an  $\ell$ - $\nu$  diagram). The first of these was obtained on July 27, 1983 during a 9.1-hour observing run, while the second was obtained during a 5-hour run on August 4, 1983. The solar p-mode ridges are clearly visible in both power spectra and are more sharply defined than they were in the intensity-only power spectrum of Figure 9. (Another version of Figure 14 in which the theoretical p-mode eigenfrequencies



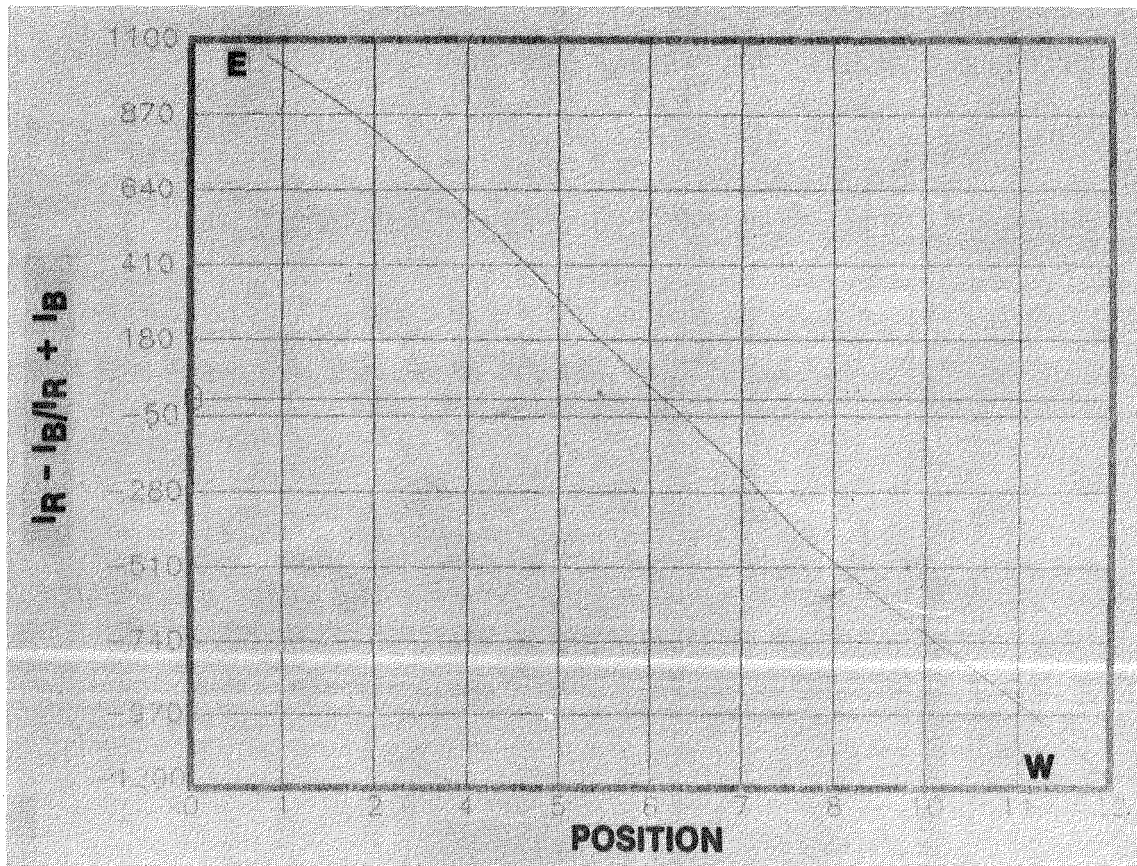


Figure 11. A calibration curve showing mean normalized intensity difference as a function of position across the solar disk. The normalized intensity differences within a small square area were averaged together and then the image was scanned across the entrance aperture of the filter and a new average was computed at each position. The slight deviations from a straight line are mainly due to the imprecise positioning employed in this preliminary test.

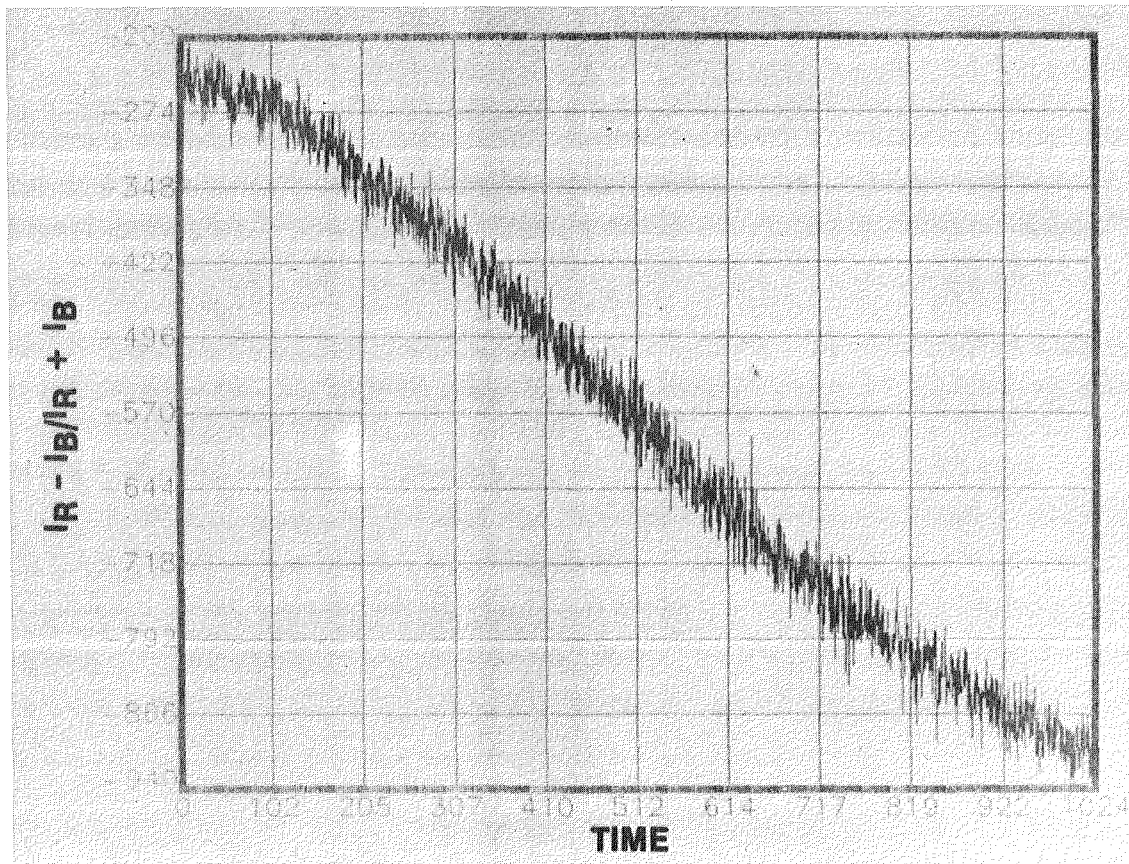


Figure 12. The normalized intensity difference was averaged over a 128 x 128-pixel square located at the center of each successive Dopplergram in a single observing run. The resulting mean values of  $I_R - I_B / I_R + I_B$  are plotted here as a function of time. Both solar oscillations and the rotation of the earth are visible here.

### COLUMN-AVERAGE VELOCITIES vs. TIME

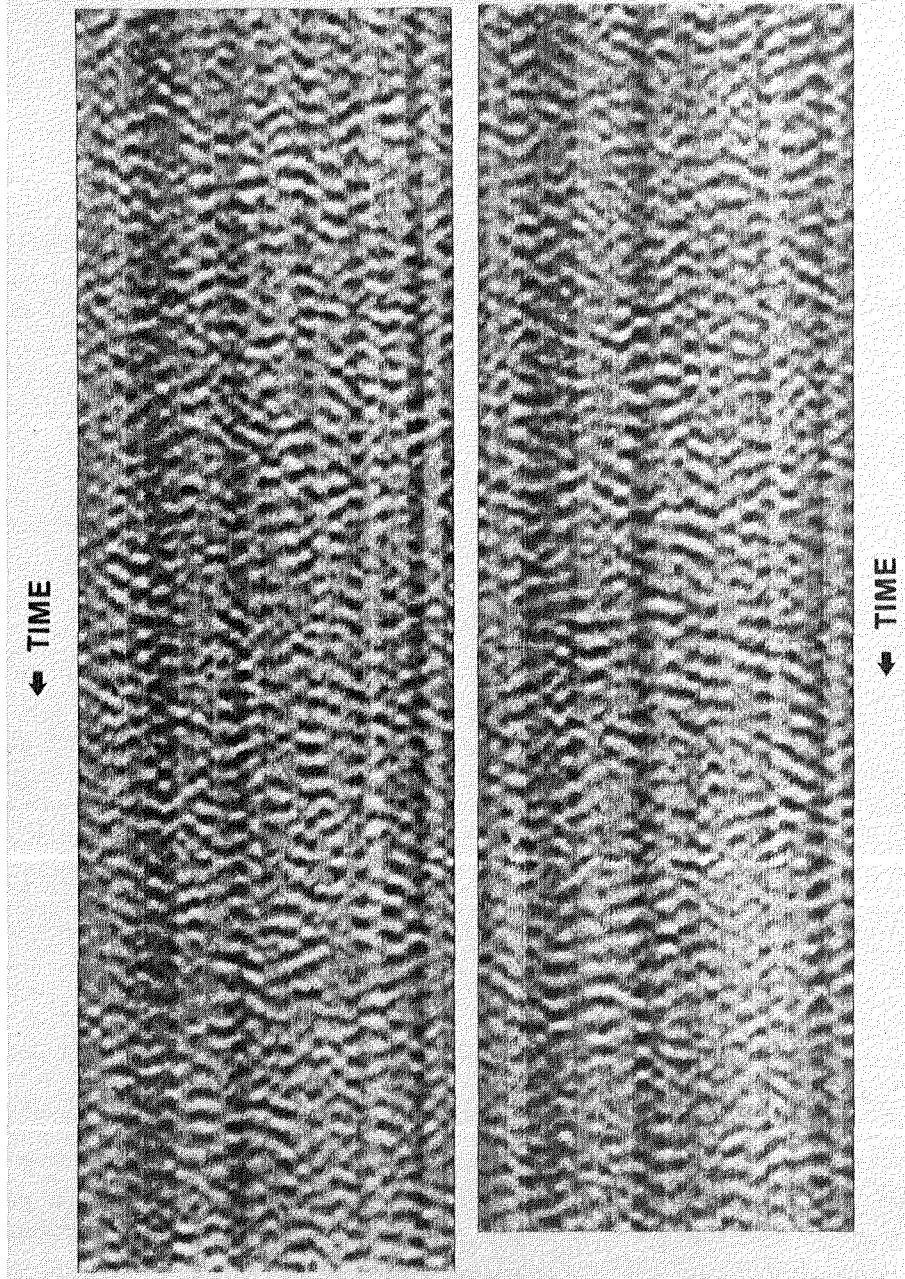


Figure 13. A space-time diagram showing temporal variability in 128 column-averaged velocities over the course of a day-long observing run. The "5-minute" oscillations are visible here as the alternating pattern of light and dark velocities. The run is broken into two halves for display and so the bottom segment is the continuation of the upper segment during the afternoon.



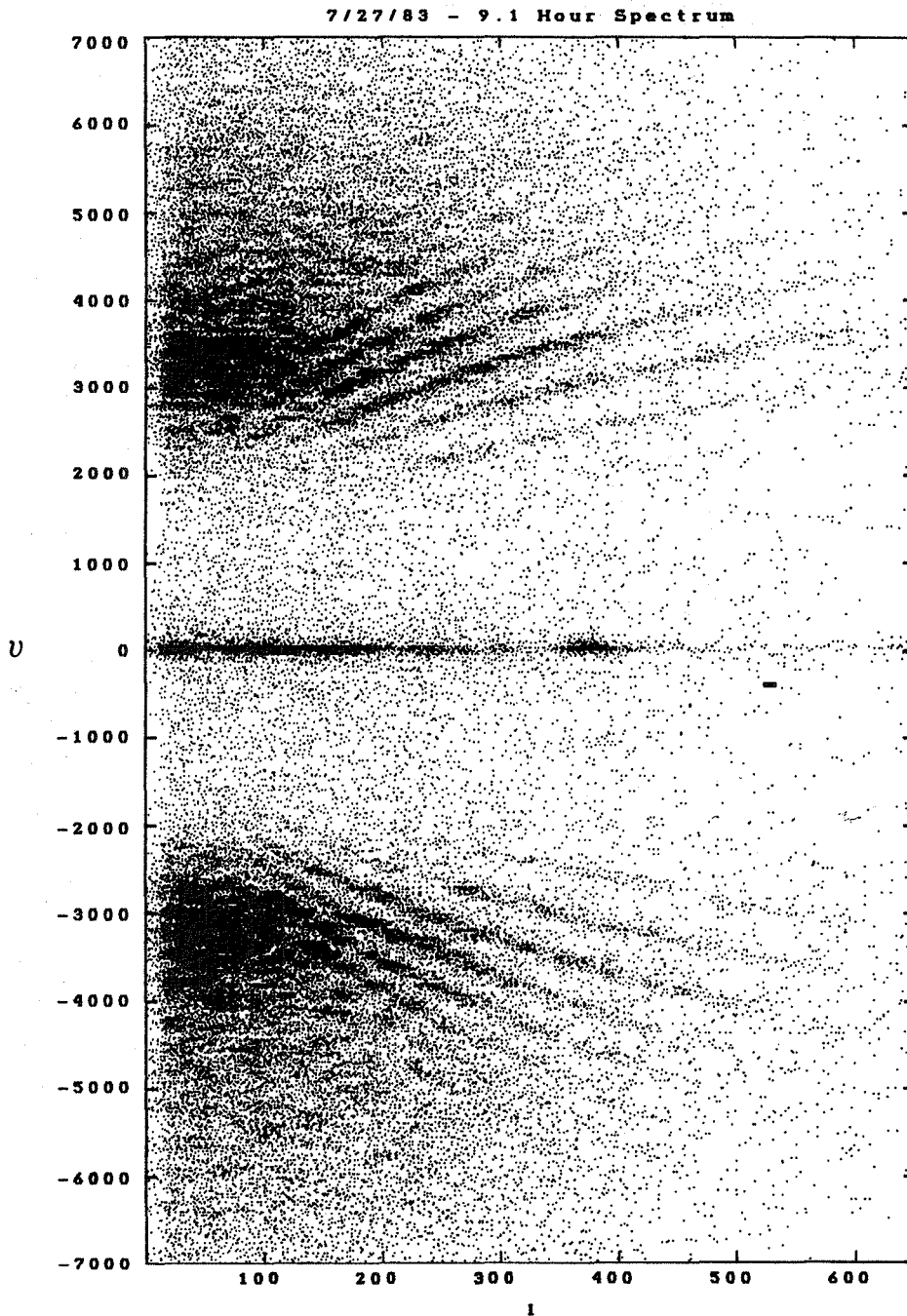


Figure 14.  $k_h$ - $\omega$  power spectrum obtained from July 27, 1983, observing run. This power spectrum was computed from the column-averaged velocities (i.e.  $I_R - I_B / I_R + I_B$  values) shown in Figure 13. Compare the quality of the  $k_h$ - $\omega$  spectrum with that of Figure 9 which was obtained using intensity information alone. The pixel size in  $\Delta k_h$  and  $\Delta \omega$  is shown as the small solid rectangle near the right center of the spectrum.

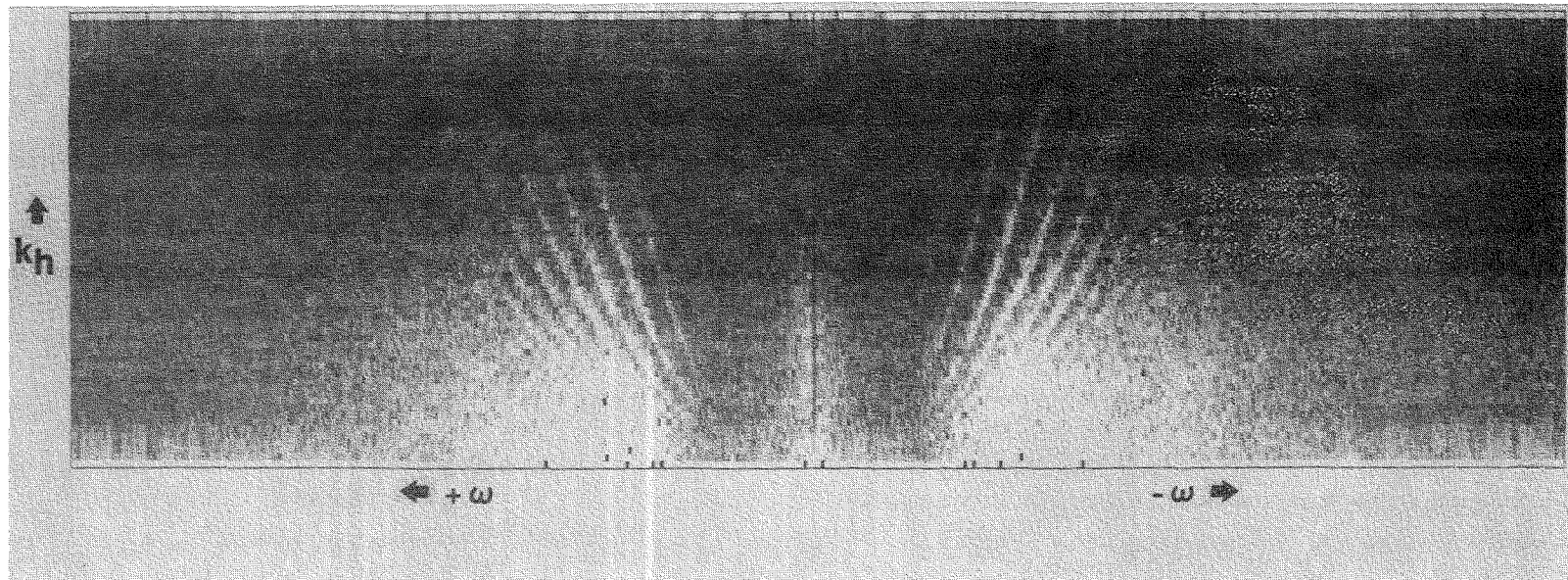


Figure 15. An average of two 5-hour Na D velocity power spectra both obtained concurrently on August 4, 1983. The telescope was moved between two nearby regions on the sun after every Dopplergram and in this way two independent time series were built up at the same time with only a small time lag between samples at the two locations. Comparison of Figures 14 and 15 shows the improvement which can be realized by improving the duty cycle of the observations.

have been superimposed on the observed ridges is contained elsewhere in these proceedings in Ulrich, et al., 1984.)

One of the reasons that the 5-hour spectrum of August 4 is almost as distinct as the 9-hour spectrum of July 27 is that on August 4 the guider was moved to one of two different locations on the sun after every Dopplergram. In this fashion two spatially separate time series of Dopplergrams were built up on this date. Each time series was analyzed separately and then the two resulting  $k_h$ - $\omega$  power spectra were averaged together to yield Figure 15. Thus, Figure 15 is indicative of the kind of gain in signal-to-noise ratio which is possible in a  $k_h$ - $\omega$  diagram due to the use of an improved duty cycle during the acquisition of the data. Had twice as many frames been available on July 27, then the spectrum of Figure 14 would be even better than that shown. Hence, it is clear that the spatial baseline on the sun should be longer than 500 arc seconds (= 128 pixels at  $3.9''$  per pixel) and that the duty cycle of the acquired frames should be as high as possible in order to yield further improvements in a  $k_h$ - $\omega$  power spectrum.

An alternative way of demonstrating the relatively high signal-to-noise ratio already present in the spectrum of Figure 14 is shown here in Figure 16. Here we show a slice at constant  $k_h$  (or constant degree,  $\ell$ ) through the power spectrum of Figure 14. The p-mode ridges stand out sharply against the relatively low noise floor which is located between the ridges.

The appearances of both Figures 14 and 15 are indicative of other solar velocity power spectra obtained with spectral lines formed at chromospheric altitudes. In particular the noise which is located between the p-mode ridges' relatively low values of  $\ell$  or  $k_h$  and which extends to both higher and lower values of  $\nu$  or  $\omega$  has been seen in numerous other solar  $k_h$ - $\omega$  power spectra obtained at Kitt Peak and Sacramento Peak using the Na D lines, CaII  $\lambda 8543$ , and H $\alpha$  (Rhodes, et al. 1981). As an example, Figure 17 contains the average of two full-day power spectra obtained at Kitt Peak during September 1981 by Duvall, Harvey, and Rhodes. The two power spectra were obtained from H $\alpha$  observations but all other chromospheric power spectra obtained at Kitt Peak looked similar. Another example for  $\lambda 8542$  is contained in Rhodes, et al. (1981). In every case of a chromospheric line the p-mode ridges have not been as distinct as in the case of photospheric-line power spectra.

##### 5. Recent Modifications to the Mt. Wilson System for Full-Disk Observing

Subsequent to the recording of the data which resulted in the power spectra of Figures 14 and 15 we have installed a new lens in the 60-foot tower telescope which has 3.5 meter focal length. With this lens we are now able to form an image of the entire sun on our CID chip where before we could only form an image of a portion of the sun on the chip. With this lens we have already obtained a few full-disk Dopplergrams, each of which has 8 arcsecond spatial resolution. Such a

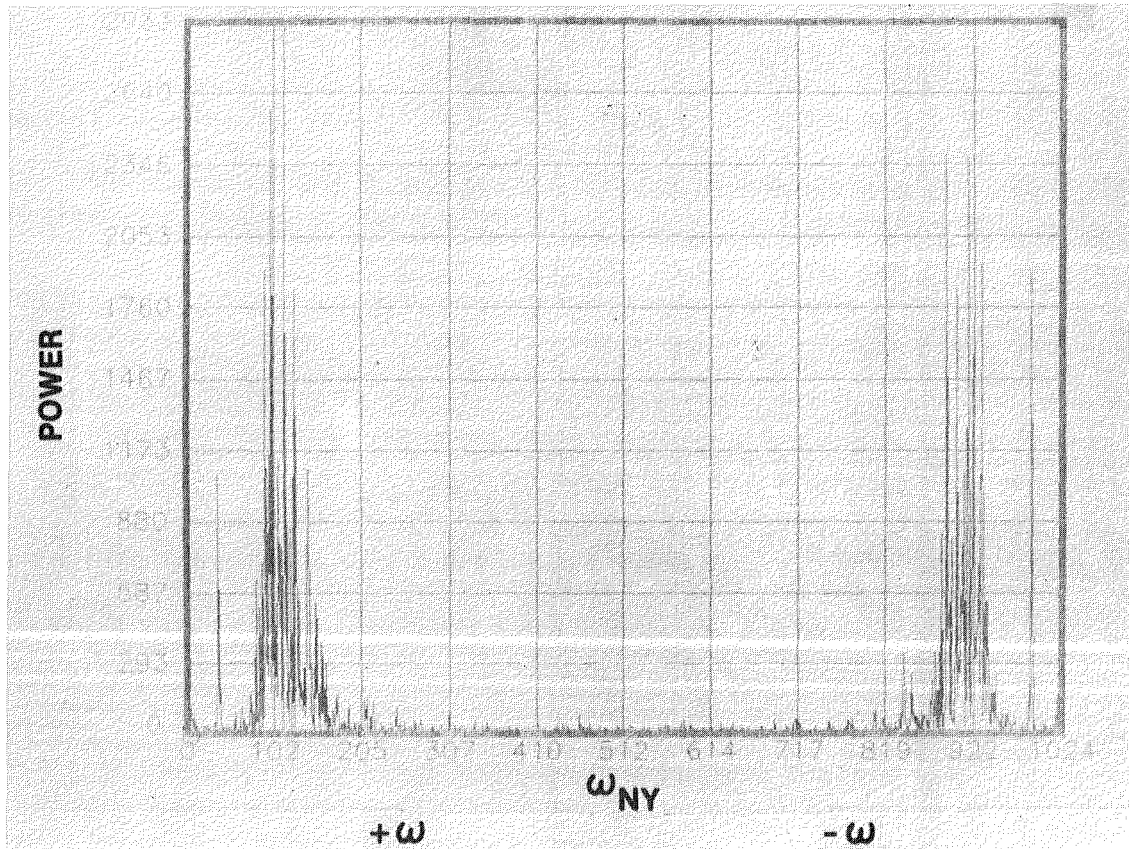


Figure 16. This is a one-dimensional slice at constant  $k_h$  through the power spectrum shown in Figure 14. The p-mode ridges are clearly visible here as the two groups of spikes present near the left and right ends of the  $\omega$ -axis. Here the negative-frequency half of the power spectrum is displayed above the Nyquist frequency,  $\omega_{NY}$ . The low noise of the filter is shown here as the low level of power away from the p-mode spikes.

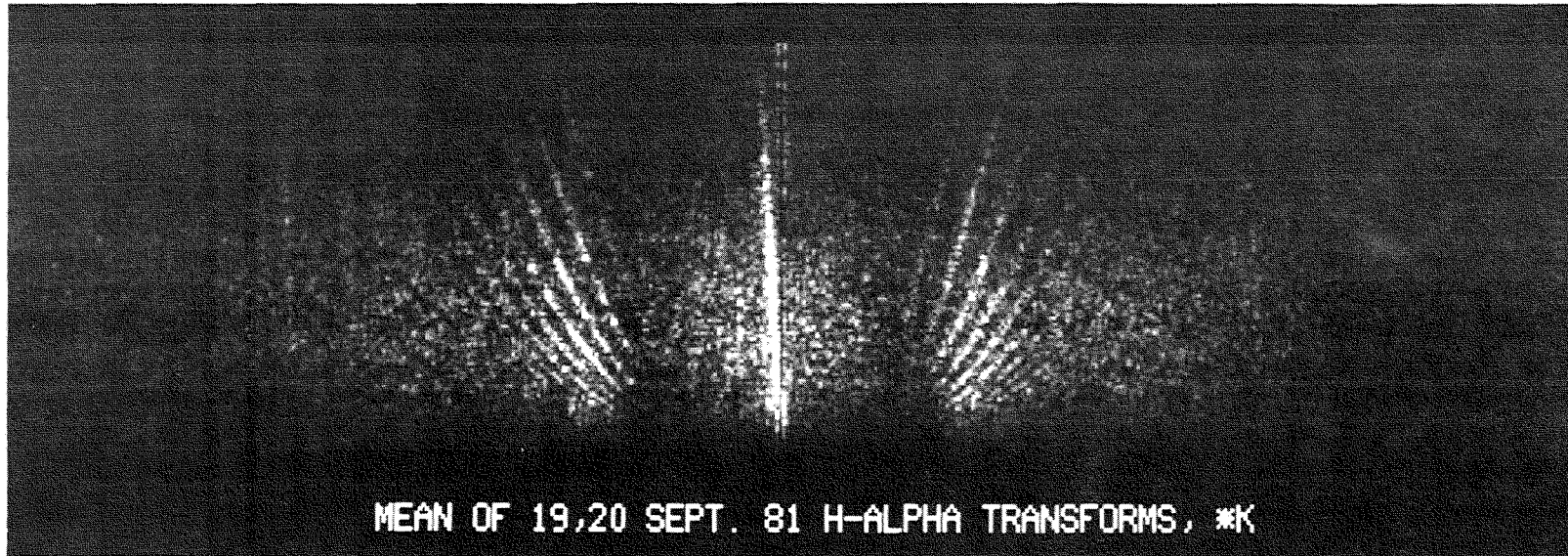


Figure 17. An average of two  $k_h$ - $\omega$  power spectra obtained at the Kitt Peak National Observatory on September 19 and 20, 1981, by T. Duvall, J. Harvey, and E. Rhodes using the McMath spectrograph. The spectral region used was near the core of  $H\alpha$  and the observing run on each of the two days was roughly 12 hours long. The randomly-distributed power at low values of  $k_h$  seen here was characteristic of all other chromospheric-level power spectra obtained at Kitt Peak.

spatial resolution is not adequate to study the high-degree solar p-modes but it is sufficient to let us study intermediate-degree modes as Duvall and Harvey (1983) recently did.

#### 6. CCD Camera Tests and Plans at Mt. Wilson

As described above full-disk Dopplergrams recorded with our existing CID camera have a spatial resolution of only  $8''$ . Consequently, we have begun experimenting with an  $800 \times 800$  pixel CCD camera supplied by the Jet Propulsion Laboratory. With this number of pixels our full-disk images will have a spatial resolution of about  $2.5''$ , or enough to study p-modes of reasonably high degree. Special interface circuitry and new cables have been constructed and we have been able to transfer pixels from the JPL camera into our array processor; however, we have not yet transferred entire CCD frames onto our disk memory. We are currently planning to install a  $1024 \times 1024$ -pixel CCD camera at Mt. Wilson during the coming summer. This camera will be fast enough to have a 1.2-second readout speed and will enable us to obtain full-disk Dopplergrams with the MOF in less than 3 seconds.

#### 7. Other Planned Magneto-Optical Filter Tests

We have also recently built two home-made magnet assemblies which will enable us to obtain 2700 Gauss each as opposed to the 1100 Gauss we are currently employing. We intend to employ these new magnet assemblies at Mt. Wilson during the summer of 1984 in efforts to see whether or not it will be possible to decrease the input power required substantially below the levels mentioned earlier. Because the higher magnetic field strengths will move the Zeeman components farther apart from each other we hope that a lower optical depth will be needed to yield the 200 mA bandpass separation than that which is currently required.

We have also acquired and plan to test a pair of potassium-filled cells. In this case we hope that the smaller dynamic range available due to the narrower line profile will still be adequate for a Langrangian point halo-type spacecraft orbit where the sun-spacecraft velocities are always very small. The principal reason for trying the potassium is that since its spectral line at  $\lambda 7699$  is a photospheric line we hope to find an even cleaner-appearing  $k_H$ - $\omega$  power spectrum than those shown earlier for sodium.

We have also obtained the first magnetograms with the two-cell Na MOF during a one-day test in August, 1983, at the Big Bear Solar Observatory and the magnetograms we obtained looked promising enough that we hope to obtain more at Mt. Wilson during the coming summer.

#### 8. Plans for Two-Station Observing with Magneto-Optical Filter

We currently are planning to collaborate with Drs. Harold Zirin and Ken Libbrecht of the Caltech Solar Astronomy program in a program of coordinated MOF observations at Mt. Wilson and at Tel Aviv University.

Initially we plan to install a second pair of MOF Na cells in Tel Aviv and then employ a video camera and video tape recorder at Tel Aviv to record the filtergrams while we are continuing to observe at Mt. Wilson with our CID and possibly also with our CCD cameras. Our initial observations are scheduled for August, 1984, with hopes to return in 1985 with enough additional electronic equipment to obtain digital images at both sites. Past experience of the Caltech group has shown that it should be possible to obtain uninterrupted oscillation observations for between 20 and 22 hours per day in this fashion.

#### 9. Pressure-Scanned Fabry-Perot Interferometer Tests

As a possible alternative to the MOF we have been studying the suitability of the Fabry-Perot Interferometer (FPI) for the acquisition of solar Dopplergrams. In particular we have been employing a normal incidence FPI as a narrow-band interference filter to image the sun in the light of the solar photospheric Fe I line at  $\lambda 5576$ . The particular FPI we are using is of the molecular adherence design in which the two etalon plates are held parallel at a fixed separation by a pair of silicon (eventually zerodur) spacers. The FPI is 5 centimeters in size with a 3 centimeter diameter working area. It has a finesse of about 30 and was manufactured in France. It was obtained for this project with the assistance of J. Blamont.

Because the two reflective plates in this interferometer are maintained (to first order) at a constant separation we have been investigating the so-called pressure-scanned method of tuning as a means of alternating between the "red" and "blue" wings of the  $\lambda 5576$  line. In our approach this has meant placing the FPI in an oven which is heated to roughly  $55^{\circ}\text{C}$  and which is filled with a gas. Changes in the pressure of the gas filling the oven result in changes in the index of refraction of the thin gap located between the etalon plates. These refractive index variations in turn tune the FPI to various places within its fixed free spectral range. Thus, for this method of operation two different gas pressures are employed which "tune" the Fabry-Perot between two bandpasses which are separated by roughly 150 mÅ and symmetrically located in both wings of the  $\lambda 5576$  line.

In order to carry out these tests at Mt. Wilson we had to first design, build, and install an entirely new optical path at the 60-foot telescope which would be large enough to accommodate the pressure-scanning oven housing the FPI. This was completed in May of 1983. We also had to modify our LSI 11 system to put out square-wave analog voltages of variable frequency. This was also done during May of 1983.

During our summer, 1983, FPI observing tests we employed an optical arrangement in which the FPI is placed in a collimated beam and in which an image of the entire solar disk is formed within the central "spot" of the interference fringe pattern at the focal plane of the CID camera. With this type of optical arrangement a systematic wavelength shift is introduced between the center of the solar disk and its limbs. One of our experimental goals was to determine the extent



to which this wavelength shift could be "calibrated out" of the resultant full-disk Dopplergrams. (Alternative optical schemes are also available in which the FPI is placed at a pupil plane and the systematic center-to-limb wavelength shift disappears; however, we have not yet converted our optical system over to such a scheme.)

The biggest problems we encountered during 1983 were in keeping the pressure constant enough during each camera exposure to minimize the refractive index-induced Dopplershifts during those exposures and in getting a high-quality image through the system. In order to deal with the pressure-scanning difficulties we have temporarily replaced the initial pressure-scanning system (which employed a Peltier cooler to change the temperature of a small drop of liquid hexane and which in turn changed the pressure within the oven) with a system of air reservoirs and electromechanical valves. With this system we can keep the pressure constant to roughly one percent during each square-wave plateau. While this will not ultimately be accurate enough we plan to at least obtain a time series of Dopplergrams with this system during the summer of 1984. Modifications to the original hexane pressure-scanning apparatus are also in progress in France and will be tried out at Mt. Wilson once they have been completed.

#### 10. Summary

Studies of the suitability of both the MOF and the FPI (at least in its pressure-scanned method of operation) for use as a compact spaceborne solar oscillation tachometer have been initiated at the Mt. Wilson Observatory. One- and two-dimensional ( $k_h$ - $\omega$ ) power spectra have thus far been obtained only with the MOF; however, the results with the MOF have been encouraging enough to warrant a continued series of tests with this device. Problems in employing the pressure-scanning approach with the FPI have been identified and alternatives devised to circumvent them. Further tests will be necessary before the FPI can be conclusively evaluated. Two-station ground-based observations with the MOF will begin during August, 1984.

#### References

- Aikens, R.S., 1980, "A Large CID for Use in Astronomy," AURA Engineering Technical Report No. 66, KPNO, Tucson, Arizona.
- Brown, T.M., 1984, "Status and Prospects for the Fourier Tachometer," these proceedings.
- Cacciani, A., and E.J. Rhodes, Jr., 1984, "The Magneto-Optical Filter: Working Principles and Recent Progress," these proceedings.
- Duvall, R.L., Jr., and Harvey, J.W., 1983, Nature, 302, 24.
- Gilman, P.A., 1979, "Dynamics of the Solar Interior and the Solar Dynamo," in Study of the Solar Cycle from Space, NASA CP-2098, p.3.
- Hale, G.E., 1908, Ap. J., 28, 315.
- Rhodes, E.J., Jr., Ulrich, R.K., Harvey, J.W., and Duvall, T.L., Jr., 1981, "The Five-Minute Oscillations: What's Left," in Solar Instrumentation: What's Next, ed. R.B. Dunn, Sacramento Peak Observatory Publication, p. 37.



Rhodes, E.J., Jr., Howard, R.J., Ulrich, R.K., and Smith, E.J., 1983, "A New System for Observing Solar Oscillations at the Mount Wilson Observatory, I. Design and Installation," in Proceedings of I.A.U. Colloquium No. 66, Solar Physics, 82, p. 245.

Ulrich, R. K., Rhodes, Jr., E. J., Cacciani, A., and Tomczyk, S., "The Effects of Seeing on Noise," these proceedings.

#### Acknowledgements

The research reported herein could not have been accomplished without the assistance of the following individuals: Thomas Andrews, Thomas Bursch, Douglas Clay, Steven Geller, Peter Kobzeff, Harry Marsh, and Thomas Thorpe, all of the Jet Propulsion Laboratory; John Boyden, Maynard Clark, and Harvey Crist of the Mount Wilson Observatory; Harland Epps, Nick Magnone, Daniel McKenna, James Wilkie, and Bradley Wood of the U.C.L.A. Astronomy Department; Wendee M. Brunish of the U.S.C. Astronomy Department, and Jeffrey Mannan of the Orbis Corporation. We also wish to acknowledge the assistance of Mr. Enzo Zante of the Italian National Laboratory for Alternative Energy (ENEA) in Frascati who helped with the fabrication of the MOF cells. We also wish to thank Jack Harvey and Tom Duvall for agreeing to the inclusion of the power spectrum obtained at Kitt Peak (Figure 17).



## THE FOURIER TACHOMETER II -- AN INSTRUMENT FOR MEASURING GLOBAL SOLAR VELOCITY FIELDS

Timothy M. Brown  
High Altitude Observatory  
National Center for Atmospheric Research  
Boulder, CO 80307

**ABSTRACT** The High Altitude Observatory and Sacramento Peak Observatory have jointly constructed a second version of the Fourier Tachometer, which is now undergoing final integration and testing. This is an interferometric instrument for measuring the Doppler shift of solar spectrum lines. The principal features and performance goals of this instrument are: simultaneous velocity observations over a 2-dimensional, 100 x 100 pixel field of view; measurement of absolute Doppler shifts with 1 m/s accuracy; noise level for moderate- $l$  oscillation modes of 1 cm/s for a 1-day observing run; flexibility and ease of use. Early (though incomplete) testing suggests that these goals should be attainable with the current instrument.

### I. Principles of Operation

The principles of operation of the Fourier Tachometer (henceforth FT) have been described elsewhere (Beckers and Brown 1978, Brown 1980), but I will review them briefly here. The incoming solar image from the feed telescope first passes through an interference filter to isolate the solar line of interest. The bandpass chosen for this filter is the result of a trade-off between systematic and random sources of error; the value we have adopted is about 3 Å. The solar image is then combined with the image of a suitable laboratory wavelength reference. These two light sources occupy different parts of the instrument's field of view, but otherwise traverse the same optical components. The beams then pass through a Michelson interferometer with a path difference of about 1 cm. The sinusoidal transmission-vs-wavelength characteristic of this element is about 0.3 Å from peak to peak, and can be shifted through at least one complete fringe by a subsequent modulating element. When one observes the output of this device integrated over wavelength, the effect of shifting the transmission pattern is to modulate the contribution of the light in each wavelength interval sinusoidally, with an amplitude proportional to the brightness at that wavelength and with a phase that depends only on the wavelength itself. If the prefilter passband is wide and smooth enough, its net contribution to the observed modulation is zero, since the different modulation phases present across its width cancel one another. However, if a localized feature

such as a solar line is present within the passband, this cancellation is incomplete. In this case, one observes a modulation of the total output intensity, with an amplitude that depends on the depth and width of the solar line, and with a phase that depends on the line's wavelength, i.e.

$$I(\phi) = I_0 \cos(\phi + \psi) ,$$

where  $\psi$  depends only on the wavelength of the solar line being observed.

Determining a solar Doppler shift thus reduces to measuring the phase of a sinusoidal variation (of known frequency) in the light output. Since the modulator imposes this sinusoidal variation simultaneously at all points in the interferometer's field of view, the device may be used to measure independent and simultaneous velocities everywhere in a 2-dimensional field, without spatial rastering. Further, by continuously monitoring a reference wavelength in a selected part of the field of view, one can measure and compensate for any inadequacies in the action of the modulator. Employing a wavelength standard also permits the solar wavelength measured to be tied directly to a known laboratory wavelength. This allows one to measure velocities relative to the telescope, rather than relative to some solar average.

Many schemes are possible for demodulating the signal. We use one of the simplest: the detector output is integrated and sampled over segments of the modulation cycle 90 degrees apart. One can then see that, relative to a time at the center of the first sampling interval,

$$I_1 - I_3 = \text{the signal's cosine component,}$$

while

$$I_2 - I_4 = \text{the signal's sine component.}$$

The desired phase is then given by

$$\psi = \arctan \frac{I_2 - I_4}{I_1 - I_3} ,$$

and the corresponding velocity signal is

$$\delta v = \frac{c \delta \lambda \psi}{2\pi \lambda} ,$$

where  $\lambda$  is the wavelength of the modulated feature, and  $\delta \lambda$  is the fringe spacing produced by the interferometer.

The FT has both advantages and disadvantages. On the positive side, it shares with certain filter-based approaches the ability to measure Doppler shifts over a full 2-dimensional field of view. Unlike most filter systems, it is inherently stable and easy to calibrate against laboratory wavelength standards at *any* visible wavelength. The detection scheme is linear, in two important senses. First, the output signal is accurately proportional to the line shift (as long as the line does not change shape), even for very large shifts from the rest wavelength. Second, averages of velocity over many pixels in the image plane depend linearly on the sine and cosine components measured in the individual pixels. This means that intensity-weighted areal averages of velocity are

independent (within the linearity of the detector itself) of any seeing-induced rearrangement of light within the area. The filters and interferometers used in the FT are high-throughput devices, which makes it possible to use detectors that have large charge wells and correspondingly low noise. For the same reason, one can reduce the time for each detector readout, in order to minimize noise from atmospheric scintillation and varying transmission. Finally, detecting a low-amplitude sine wave is a chore well-suited to current detector technology, since one need not worry about pixel-to-pixel variations in the gain or dark current of the detector, or about time variations in these quantities.

There are two principal disadvantages inherent in the FT approach. The more serious is that, for an acceptably low level of sensitivity to the characteristics of the line-isolating prefilter, the bandpass of this prefilter must be rather large compared to the width of the solar line. This causes a large, unvarying background to be superposed on the modulated signal one wishes to measure. The background contributes nothing to the signal, but does add to the noise. Thus, the FT suffers from noise levels that are poor relative to the best theoretically achievable by some other means. However, as we will see shortly, the noise level one should be able to reach with the FT compares favorably with the best current practice, and is quite good enough to provide new information about the Sun. The lesser problem with the FT (at least for most applications) is that the detection scheme explicitly ignores all line shape information, except for a strength parameter that is similar to a measurement of equivalent width. Experiments that require detailed line shape information, such as efforts to detect atmospheric gravity waves, will probably not profit from the FT.

## II. Fourier Tachometer II Features and Capabilities

The new FT is being installed and tested at the Big Dome of Sacramento Peak Observatory (now part of the National Solar Observatory). It is intended to be a user instrument, and for this reason it has been designed with enough flexibility to address a large range of observational problems. Most of the features outlined below are already functional, though some are not. These are explicitly noted, and in any case are expected to become available within the next year.

The interferometer used in the new FT is a polarizing, wide-field Michelson that has been described in detail by Evans (1980). This device behaves like a single Lyot filter element, retarding one polarization component with respect to the other by roughly 1 cm. However, by comparison with a Lyot element it admits a very wide field. With proper construction, the 2nd order variation of retardation with field angle can be made to vanish, leaving the 4th order term as the most important. Moreover, by proper choice of glasses in the two legs, the thermal variation of retardance can be made quite small. Considering the probable errors in published values for the variation of the glasses' refractive index with temperature, the thermal variation of apparent line position will probably be less than 700 m/s per degree C. Modulation of the fringe position is accomplished by rotating a half-wave plate following the interferometer.

Since the polarizing beamsplitting coatings available can be optimized only over a limited wavelength range, and since the best interferometer path difference is inversely proportional to the width of the line one wishes to measure, we will eventually have three different interferometers. Only one of these is yet available; it is designed for weak solar lines between 6300 and 6800 Å. The second will also have a path difference appropriate to weak lines, but will be coated to operate between 5300 and 5800 Å. Finally, there will be an interferometer with a significantly smaller path difference, optimized and coated to allow observation of the Na D lines and the Ca II infrared triplet.

By choosing among prefilters and interferometers, one will be able to observe up to 6 solar lines in quick succession, and in any order. Typically about 1 s will be required for the computer to move from one filter to another. We plan to have a number of standard filters available, including ones for Fe I 5434, 5576, 6302, 6678, Na D 5896, and Ca II 8542. Observers wishing to use some other line will be able to do so, simply by providing an appropriate prefilter. The only filter currently available is for Fe I 6678.

A variety of reference wavelength sources are available for use with the FT. For calibration of field-dependent errors in the velocity signal, one may use either an image of the objective lens or the light from a Zeeman-stabilized He-Ne laser. The objective image has the advantage of providing an absorption feature of nearly the same width as the line actually being observed, and at the same mean wavelength. It has the disadvantage of not being tied to a particular laboratory wavelength; as a result it shows variations in its center wavelength on all time scales, arising from motions of the Earth, variations in the atmospheric extinction seen across the solar disk, and various solar processes. The stabilized laser is of the type described by Baer and Hall (1980). Its stability on time scales of a day or so is probably a few times  $10^{-10}$ , while for timescales of a year the stability is likely to suffer from variations in the gas pressure in the laser tube. These variations should cause shifts in the output wavelength of a few times  $10^{-9}$ , corresponding to velocity errors of about 1 m/s. The principal advantage of the laser as a calibration source is thus its ability to provide a bright source of light at a known wavelength in the lab. Its disadvantages are its very narrow line width, and the large wavelength interval between the 6328 Å laser line and most of the interesting solar lines.

For continuous monitoring of the modulator performance, and of drifts in the interferometer, we have a neon discharge lamp with filters to select neon lines near a wide selection of solar lines. These lines are expected to have a relative wavelength stability of  $10^{-8}$  or so over the life of the discharge tube, and to vary little from day to day. Filters for the neon source may be selected and changed under computer control.

The solar image for the FT is provided by the 40 cm coronagraph at the Big Dome. The feed optics provide three image scales, corresponding roughly to 20, 5, and 1.2 arcsec per pixel at the detector. At the 20 arcsec per pixel scale, the entire solar image fits on the detector at one time; at the 1.2 arcsec per pixel scale, one can see an area on the Sun roughly the size of a typical active

region. The selection of image size is done manually, and about a minute is required to change image scales. Currently, we have computer control of the image rotation on the detector. Shortly, however, we expect to add a system that will allow complete control of the position of the image center on the solar disk.

All of the computer-controlled functions of the instrument (filter selection, image rotation, light source selection, timing and data taking) are controlled by a special-purpose instrument programming language (the Hands-On Operating and Programming Language, or HOOPLA). Besides providing a convenient and flexible means of operating the instrument, this language provides a powerful set of real-time diagnostic displays, and supervises all data transfers to the Big Dome's PE 3220 mainframe computer. Software available on the 3220 provides for both short and long term data storage, and for a variety of analysis and display functions.

A number of sources of error are likely to affect the velocity measurements done with the FT. Until testing is complete, it will not be clear which of these is causing the most trouble. However, simple estimates suggest the following may be of some importance.

If the solar line being observed lies on a sloping part of the prefilter transmission curve, its apparent wavelength will be shifted towards the direction of decreasing transmission (for an absorption line). A similar shift results from transmission variations in other optical elements, caused for example by fringing effects. The size of this line shift is proportional to the width of the line being observed, so the laser and neon reference lines are much less affected than the solar lines, and cannot be used to remove the problem. We expect that this will be the major source of error in the average solar velocity as measured by the FT; with plausible temperature control on the prefilters, the noise from this source should be a few m/s, with most of its variation at periods longer than about ten minutes.

Readout and shot noise in the Reticon appear to be almost completely random, both in space and in time. The observed ratio of peak signal to rms noise is about 2000:1. With an assumed modulation fraction of 0.02 and an integration time of 60 s, this leads to an estimated random velocity noise of less than 5 m/s for each pixel. Spatial spectra of the noise show that about 1% of it is correlated among the pixels of a given line, and about 0.1% arises from an as yet unidentified source of harmonic noise. These facts suggest that the noise will decrease as desired when spatial filtering operations are performed on the data, giving a readout noise contribution to full-disk average velocity measurements of only a few cm/s in a 60 s integration.

Besides the two just mentioned, several other noise sources are expected to be less important. These include uneven rotation of the modulator waveplate, a number of atmospheric effects, variation of the interferometer field-dependent errors with temperature and time, and instability of the reference wavelength

sources. Of course, only complete testing will reveal whether the importance of these effects is as small as calculations indicate, and whether some other significant source of noise has been overlooked. However, assuming that the calculations are approximately correct, one can assemble a list of reasonable performance goals for the FT. These take the form of the minimum amplitude narrow band signal that one might detect in an 8-hr observing run, given as a function of the frequency  $\nu$  of the signal and of its total angular degree  $l$ . These values are given in the accompanying table, in units of cm/s. For  $l = 0$ , the principal limit to detection will likely be the variations of filter passband wavelength. At frequencies below 1 cycle per day, the errors from this source will probably be 2 m/s or so, improving little with lengthening observing time. At higher temporal frequencies, it seems reasonable to take the noise power as proportional to  $1/\nu$ , with the ultimate goal of reducing the rms noise integrated over frequency to below 1 m/s. For  $l$  values above 10, or for  $\nu$  larger than 1 mHz, one expects the dominant source of noise to be the spatially and temporally white detector readout noise. A plausible goal is thus to detect oscillations in this range with amplitudes of about 1 cm/s. If the noise proves to be truly white, one might do somewhat better. For the intermediate range of  $l$  values between 1 and 10, and for  $\nu$  below about 1 mHz, one can expect something between the two previous cases. Exactly how well one should expect to do is not clear, however.

REASONABLE PERFORMANCE GOALS

	1-10	3-8	1	1
FREQ (mHz)	<1	8-40	?	1
	0	200	?	1
		0	1-10	>10
		$l$		

In summary, the Fourier Tachometer now being installed at SPO will be a user instrument with many capabilities. It will provide a 2-dimensional field of view, a noise level competitive with the best current velocity instrumentation, great temporal stability, and access to a number of solar spectrum lines and image scales. Its first applications will undoubtedly be to solar oscillations, where it may make valuable contributions.

Two institutions and many people have been involved in the construction of the FT II. Sacramento Peak Observatory took principal responsibility for the optical and mechanical components, while the High Altitude Observatory was responsible for most of the electronics, the stable laser, and the software for instrument operation and data reduction. While it is not possible to list



everyone involved, a few people deserve special mention. At SPO the prime mover has been Jack Evans, assisted by George Streander, Al Healy and Wayne Jones. At HAO, most of the credit goes to David Elmore, Greg Muir, Terry Leach and Howard Hull. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

#### REFERENCES

- Baer, T. and Hall, J. 1980, in *Solar Instrumentation -- What's Next?*, R.B. Dunn, ed., Sunspot, p. 142, National Solar Observatory Publication.
- Beckers, J.M. and Brown, T.M. 1978, paper presented at JOSO Workshop on Future Solar Optical Observations, Florence.
- Brown, T.M., 1980, in *Solar Instrumentation -- What's Next?*, R.B. Dunn, ed., Sunspot, p. 151, National Solar Observatory Publication.
- Evans, J.W. 1980, in *Solar Instrumentation -- What's Next?*, R.B. Dunn, ed., Sunspot, p. 155, National Solar Observatory Publication.



# OBSERVATIONS OF INTERMEDIATE-DEGREE SOLAR OSCILLATIONS

J. W. Harvey  
Kitt Peak National Observatory\*

T. L. Duvall, Jr.  
Laboratory for Astronomy and Solar Physics,  
NASA/Goddard Space Flight Center

**Abstract.** A progress report on observations of intermediate degree oscillations is presented. We list frequencies of zonal p-mode oscillations with amplitudes in excess of  $\sim 2 \text{ cm s}^{-1}$ . These frequencies show systematic disagreement with recent theoretical calculations. We compare the frequencies with asymptotic formula estimates. Small scatter is obtained for low degree modes but large scatter at large degree. A first look at sectoral harmonic observations shows that magnetic active regions provide a major signal at low frequencies.

## 1. Introduction.

In 1980 we undertook a series of observations of solar oscillations of intermediate spherical harmonic degree,  $l$ . The initial motivation for these observations was to obtain frequencies of p-modes in the previously unobserved degree range from  $\sim 5$  to  $\sim 100$ . Further motivation was to try to observe rotational splitting of these modes, to observe g-modes of low and intermediate degree and to explore the latitudinal variation of oscillation patterns.

This report follows two previous reports (Duvall and Harvey, 1983; Pomerantz *et al.*, 1982) on aspects of this observing program.

## 2. Observations.

We have observed solar oscillations by means of the Doppler shifts they produce in photospheric spectrum lines and also the intensity fluctuations they produce in spectrum lines formed near the atmospheric temperature minimum. The Doppler observations were made at Kitt Peak using a large grating spectrograph equipped with a two-dimensional diode array camera at the spectrograph focal plane. The techniques were developed from earlier measurements of high-degree oscillations made in collaboration with E. Rhodes, Jr. and are partly described by Rhodes *et al.* (1981) and Duvall and Harvey (1983). Some intensity observations were made at Kitt Peak but the bulk of our intensity measurements were made in collaboration with M. Pomerantz from the

\*Kitt Peak National Observatory is operated by the Association of Universities for Research in Astronomy, Inc. under contract with the National Science Foundation.

geographic South Pole. The latter observations are described by Harvey *et al.* (1982) and Pomerantz *et al.* (1982).

To simplify the data handling without compromising our scientific goals too much, we restricted the angular scale of the observations to about 10 arc seconds. This restriction leads to considerable attenuation of modes of degree larger than  $\sim 200$ . We observed the full image of the sun but, in the case of the Doppler measurements, the image was optically averaged in one dimension before entering the spectrograph slit. To isolate zonal harmonics the slit is aligned parallel to the rotation axis of the sun, defined by the Carrington rotation axis, and the image is averaged in the perpendicular direction. To isolate sectoral harmonics the geometry is rotated 90 degrees. The degree of isolation of zonal harmonics offered by this technique is fairly good since zonal harmonics project onto the plane of the sky as nearly one-dimensional functions that are not distorted much by our optical averaging. The isolation of sectoral harmonics is not as good because of the curvature of the meridians as seen in projection. In practice, the zonal observations have a degree resolution of about 3 and the sectoral observations about 4. The full disk intensity measurements should isolate degree with a resolution of about 2. The same numbers characterize our ability to isolate modes of different angular order,  $m$ .

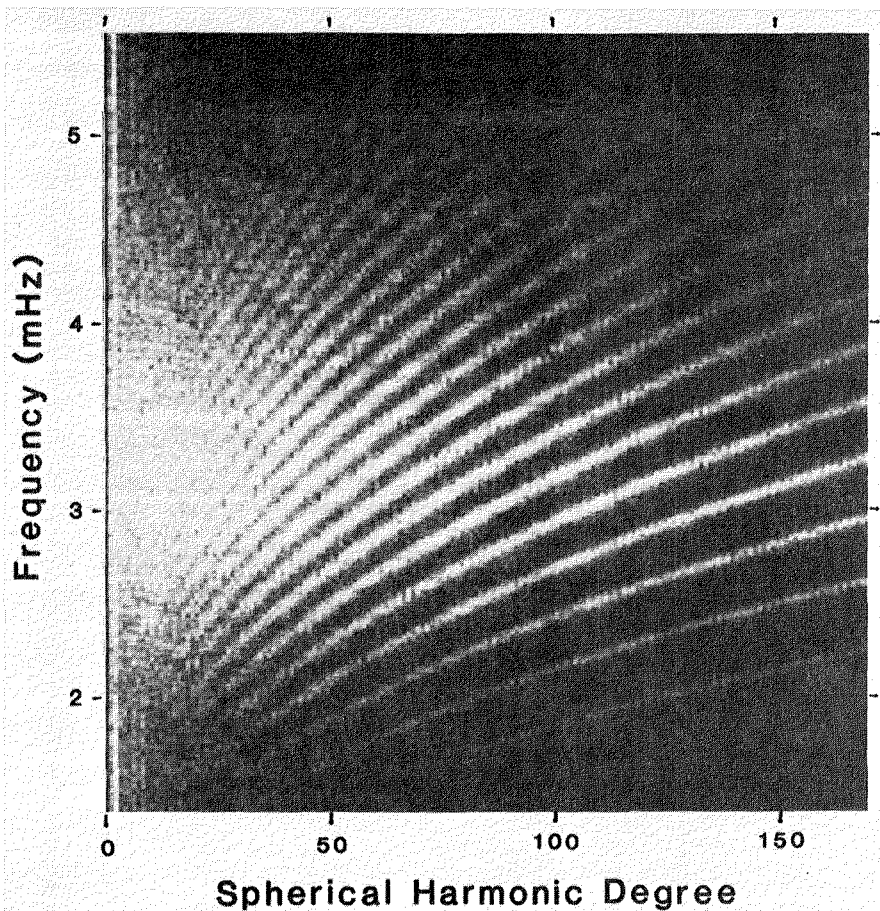
In the reduction phase, the raw observations are treated by various numerical filters to reduce the effects of instrumental noise. Each cleaned-up time sample is transformed into a series of numbers that is the product of the image times a set of slightly modified, projected spherical harmonic functions. Although we have experimented with several techniques for filling the data gaps caused by clouds and night, our results to date are based on simple Fourier transformation of unwindowed and unapodized, gapped data strings in time. Thus a significant amount of oscillatory power appears in spectral sidelobes. Table 1 is a summary of the best observing sequences.

Table 1. Observation summary.

type	angular order	dates	coverage (%)	duration <sup>-1</sup> ( $\mu\text{Hz}$ )
Doppler	zonal	1981 June 4-7	40	3.8
		1982 September 13-18	34	2.0
	sectoral	1983 May 10-26	47	0.7
Intensity	all	1981 December 24-25	92	5.2
		1981 December 5-		
		1982 January 10	20	0.3

### 3. Results

Results from the June 1981 observations have been given by Duvall and Harvey (1983). We present here preliminary measurements of the September 1982 zonal observations that are substantially better than the June 1981 data. Figure 1 shows a spectrum of the September 1982 observations. The noise level in quiet parts of the spectrum is  $\sim 1 \text{ cm s}^{-1}$  amplitude in a frequency and degree resolution element. With this noise level one sees p-modes with radial orders,  $n$ , from 1 to 25 as ridge structures. The strongest modes have amplitudes of  $\sim 10 \text{ cm s}^{-1}$  so the signal-to-noise ratio is only about 10. As the noise is instrumental, there is room for much improvement in solar oscillation spectroscopy!



**Figure 1.** Spectrum of zonal p-modes observed in September 1982. The frequency range shown is 1383 to 5550  $\mu\text{Hz}$  and degrees from 0 to 170 are plotted. The faint ridge in the lower right corner is radial order one.

To determine the frequencies of individual modes in the presence of imperfect degree and frequency resolution, we use an *echelle* method of displaying the data. Briefly, each ridge is displayed on an interactive graphics terminal with the base frequency of each successive value of degree shifted so that the ridge appears to be a straight vertical line. This is done by fitting a spline function to either the centroid of the ridge or individual modes whenever the latter are clearly identifiable. This procedure yields frequencies that are good to a few  $\mu\text{Hz}$ . The frequencies are then improved by manual examination of a printed version of the spectrum. Table 2 is a list of frequencies from the September 1982 data determined in this manner. The weakest identified modes have amplitudes of  $\sim 2 \text{ cm s}^{-1}$ .

Compared with our previously published list of frequencies, the present list contains more modes and should be more accurate, i.e. about  $\pm 1 \mu\text{Hz}$ . There is a typographical error in our previous list. The  $l=4, n=20$  mode should have been printed as 3137  $\mu\text{Hz}$ . With that correction, we find that the present frequencies are 0.41  $\mu\text{Hz}$  larger than the previous values, with a standard deviation of the differences of  $\pm 1.13 \mu\text{Hz}$ . There is a considerable amount of fine structure in the spectrum of many of the modes which limits the attainable frequency precision.



Table 3 is a list of frequencies of the spline functions that were fit to each ridge of constant radial order. These frequencies may be in error by as much as  $\pm 10 \mu\text{Hz}$ , especially at high degree.

Table 3. Frequencies of spline functions ( $\mu\text{Hz}$ ).

$l$	20	40	60	80	100	120	140	160	180	200
$n$										
1							1722	1810	1892	1967
2				1694	1849	1978	2091	2198	2300	2400
3			1781	1969	2148	2299	2428	2549	2664	2770
4		1741	2009	2234	2424	2591	2742	2884	3014	3125
5		1941	2235	2472	2679	2870	3042	3196	3340	3483
6		2134	2441	2700	2928	3133	3318	3490	3650	3798
7		2317	2644	2923	3168	3388	3590	3782		
8	2050	2497	2845	3134	3398	3640	3863	4058		
9	2204	2675	3035	3346	3625	3876	4103			
10	2352	2849	3227	3552	3846	4115	4362			
11	2498	3020	3415	3756	4064	4342	4597			
12	2645	3189	3598	3957	4277	4560				
13	2791	3355	3793	4152	4487	4791				
14	2937	3515	3962	4340	4689					
15	3081	3683	4152	4536						
16	3224	3838	4318	4723						
17	3368	4002	4521	4939						
18	3514	4152	4676							
19	3655	4302	4854	5306						
20	3799	4448	5027							
21	3943	4600	5173							
22	4091	4751								

We have used the frequencies in Table 2 to fit the well-known asymptotic approximation for  $n \gg l$ . The equation is given in a nonstandard but more practical form here as,

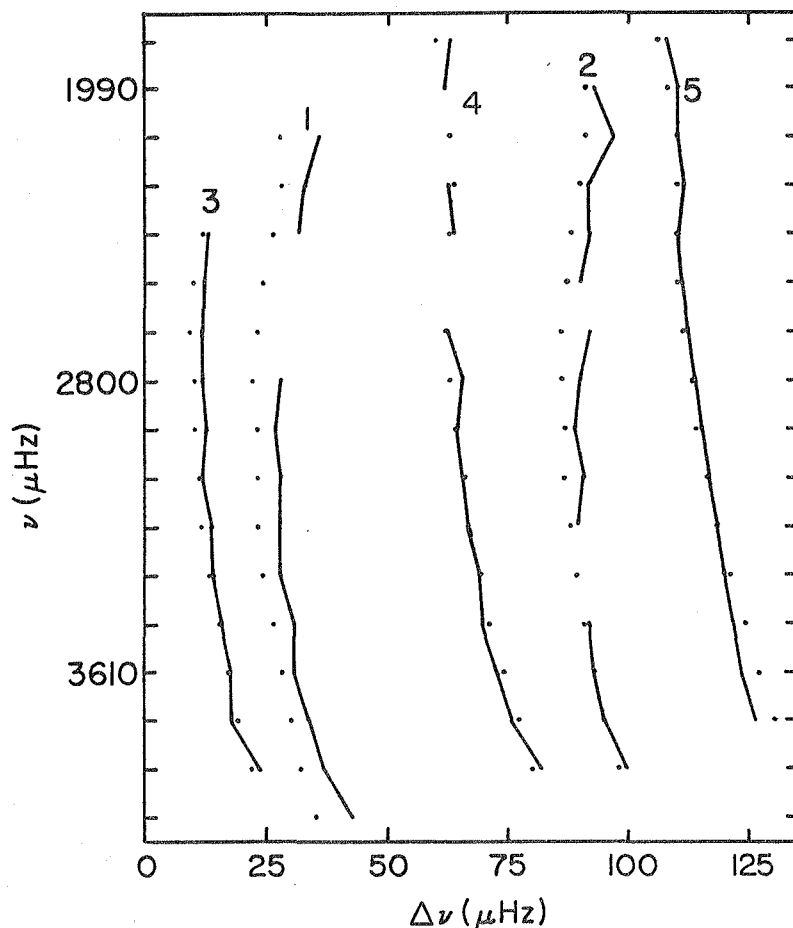
$$\nu \approx \frac{\nu_0}{2} \left[ \left( n + \frac{l}{2} + \epsilon \right) + \left\{ \left( n + \frac{l}{2} + \epsilon \right)^2 - 4A[l(l+1) + \delta] \right\}^{1/2} \right]$$

For the ranges  $l$  (1-5) and  $n$  (11-27), an unweighted least squares fit gives  $\nu_0 = 138.66 \mu\text{Hz}$ ,  $A = +0.2752$ ,  $\epsilon = +0.2553$ , and  $\delta = -49.99$ . The residuals in the range were about  $\pm 2 \mu\text{Hz}$  but were rapidly increasing at the limits of the range.

We used the frequencies in Table 3 (and for every intermediate value of  $l$  as well) to fit various asymptotic formulae for frequencies of high degree. None of the formulae we tried worked very well. Residuals were quite large. Our conclusion was that even  $l=200$  is not a "high" value in the asymptotic sense. A new formula, based on the cavity model of p-modes (Duvall, 1982), has been derived that fits all the frequencies for degrees greater than 0 and will be described elsewhere (Harvey and Duvall, 1984).

It is interesting to compare our measurements with recent theoretical calculations by Shibahashi *et al.* (1983). This is done in the *echelle* plot of Figure 2. The observed

frequencies are systematically greater than theory for degrees 1,2 and 3. Agreement at degree 4 is quite good. A systematic discrepancy that reverses sense at high and low frequency appears at degree 5.



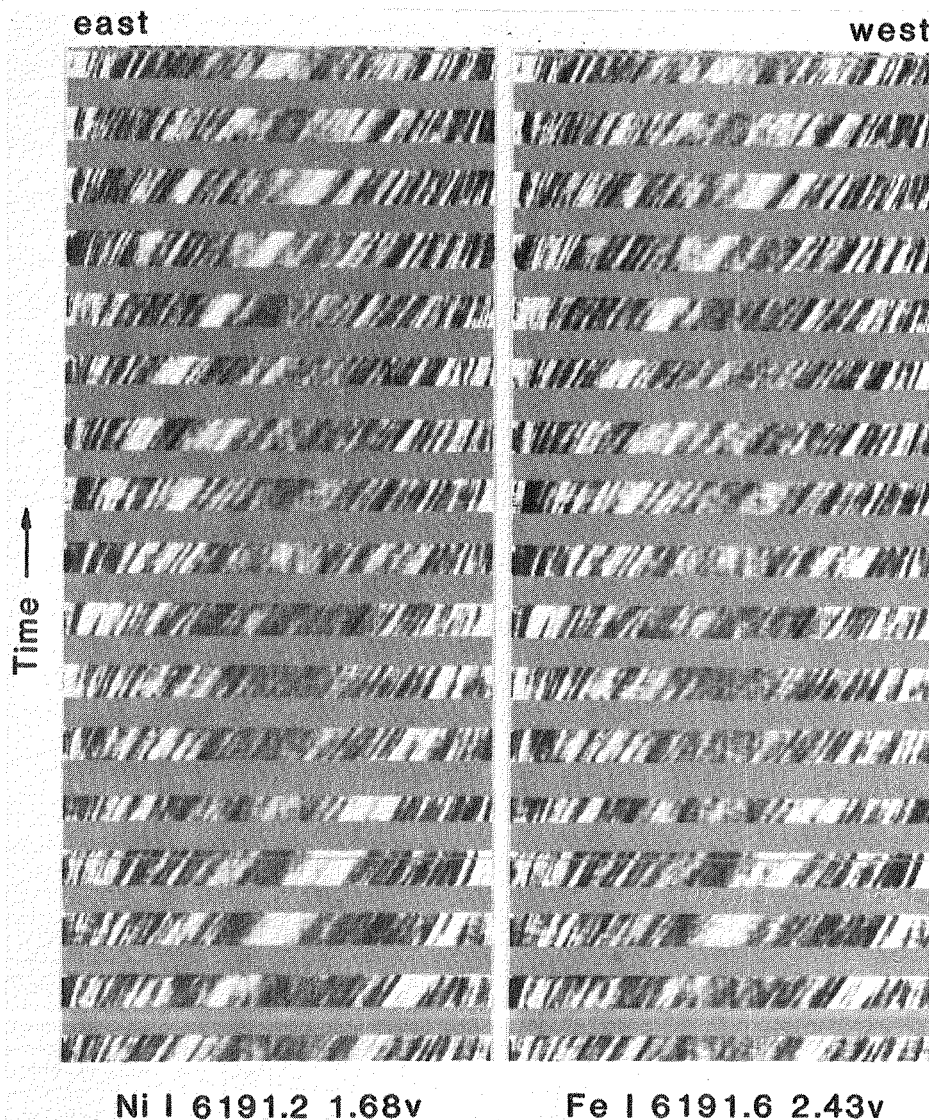
**Figure 2.** Comparison of observed frequencies from September 1982 (—) with those computed (•) by Shibahashi *et al.* (1983) for their model 1. (There is an error in their tabulation for  $l=4$ ,  $n=16$ .) This *echelle* plot has a step frequency of  $135 \mu\text{Hz}$ .

The modes are significantly narrower in frequency width at low frequency than at high frequency. This has been noted previously (e.g. Grec *et al.*, 1983). Presumably this is due to a relatively short lifetime of high frequency modes in accord with theoretical predictions (cf. Christensen-Dalsgaard and Frandsen, 1983).

At low frequencies, the faint structure that appeared in our June 1981 spectrum at a frequency around  $450 \mu\text{Hz}$  does not appear in the September 1982 observations. The cause of this artifact is not clear to us. We have looked for g-modes in the September 1982 data but our poor frequency resolution and noise from harmonics of one day prevent us from clearly seeing any g-modes. Any such modes would have to have amplitudes less than  $10 \text{ cm s}^{-1}$  or they would be clearly visible in the September 1982 spectrum.



At this time only very preliminary results from the May 1983 sectoral observations are available. Figure 3 shows a time and space plot of the Doppler shifts obtained using two different spectrum lines observed simultaneously during this observing sequence. Comparison with daily magnetogram observations shows that the light features that rotate across the disk are active regions. Thus, a systematic red-shift of  $\sim 10 \text{ m s}^{-1}$  is associated with magnetic activity in our observations. The two spectrum lines respond somewhat differently to the active regions in the sense that the higher excitation potential line shows a smaller red-shift. This suggests that the active region signal may be due in part to temperature sensitivity and that a different spectrum line might show a much weaker response to active regions. In the May 1983 data, it is clear that the major signal at low frequencies is due to active regions and not to instrumental effects.



**Figure 3.** Seventeen consecutive days of observations of the Doppler shifts of two spectrum lines from the east to west limbs. Each measurement is a north-south integration across the disk in an effort to isolate sectoral modes. The Doppler shifts are shown as red (light) and blue (dark). Active regions are seen as light streaks rotating from east to west.

Work is in progress to measure frequency splittings of the sectoral p-modes from degree 1 to 200 with a precision of  $\sim 0.1 \mu\text{Hz}$  using both the May 1983 and South Pole data. Rotational splitting in the May 1983 data is clearly visible from degrees 1-200 and will be presented elsewhere (Duvall and Harvey, 1984). The South Pole data will also be used to probe latitudinal effects from measurements of the frequencies of tesseral modes.

#### 4. Acknowledgements

We are indebted to J. Christensen-Dalsgaard, P. Goode, D. Gough, J. Leibacher, E. Rhodes, Jr., R. Ulrich and J. Toomre for helpful discussions. This work was done while T.L.D. was a visiting astronomer at Kitt Peak.

#### 5. References

- Christensen-Dalsgaard, J. and Frandsen, S.: 1983, *Solar Phys.* **82**, 165.  
Duvall, T. L., Jr.: 1982, *Nature* **300**, 242.  
Duvall, T. L., Jr. and Harvey, J. W.: 1983, *Nature* **301**, 24.  
Duvall, T. L., Jr. and Harvey, J. W.: 1984, *Nature*, in press.  
Grec, G., Fossat, E. and Pomerantz, M. A.: 1983, *Solar Physics* **82**, 55.  
Harvey, J. W. and Duvall, T. L., Jr.: 1984, *Solar Phys.*, in press.  
Harvey, J., Pomerantz, M. and Duvall, T., Jr.: 1982, *Sky and Telescope* **64**, 520.  
Pomerantz, M. A., Harvey, J. W. and Duvall, T., Jr.: 1982, *Antarctic J. of U. S.* **17**, No. 5, 232.  
Rhodes, E. J., Jr., Ulrich, R. K., Harvey, J. W. and Duvall, T. L., Jr.: 1981, in *Solar Instrumentation: What's Next?* (ed. Dunn, R.), 37, Sacramento Peak Observatory, Sunspot.  
Shibahashi, H., Noels, A. and Gabriel, M.: 1983, *Astron. and Astrophys.* **123**, 283.

## DETECTION OF SOLAR GRAVITY MODE OSCILLATIONS

Philip H. Scherrer  
Center for Space Science and Astrophysics  
Stanford University, ERL  
Stanford, California 94305 - U.S.A.

### *Abstract*

An analysis of solar velocity data obtained at the Stanford Solar Observatory has shown the existence of solar global oscillations (Delache and Scherrer, *Nature*, in press). The oscillations are in the range 45 to 105  $\mu\text{Hz}$  (160 to 370 minutes) and are interpreted as internal gravity modes of degree  $l=1$  and  $l=2$ .

### *The Data*

Observations of the global solar velocity field have been recorded at the Stanford Solar Observatory since 1976. These observations have recently been examined for long period oscillations at periods other than 160 minutes. Delache and Scherrer (1983) have reported the results of this analysis. This report will review that work and expand on the methods used to find significant peaks in the power spectrum in the presence of noise and gaps in the data.

Scherrer and Wilcox (1983) have described the observing procedures and preliminary data reduction applied to the observations. The observations are differential measures of the line-of-sight velocity of the solar surface. They are made by comparing the average doppler shift from the center of the solar disk with the doppler shift from a concentric annulus. Since it appeared that interesting features were present in the low frequency part of the spectrum, the computation was extended down to  $\nu=0$ . Since only the low frequency part of the spectrum is examined here, the data was averaged into 5-minute intervals and normalized within each day as in the previous analysis.

First the Fourier transform of the 4 years of Stanford data (1977-1980) was computed using a standard fast Fourier transform code (FFT). It was found that the largest power is in the range 45  $\mu\text{Hz}$  - 105  $\mu\text{Hz}$  (about 160 to 370 minutes). The resulting power spectrum shows a number of sets of lines with separations corresponding to day side bands (i.e. 11.57  $\mu\text{Hz}$ ) which obviously come from the nightly gaps in the data. Comparing the 4-year spectrum with individual yearly spectra, it was found that the strongest lines were more prominent in 1979 than in the other years. This may be due to the relatively clear skies and the distribution of observing times in 1979. As a first step, the analysis was restricted to that year. Figure 1a shows the 1979 spectrum. The observations were begun on 7 April and continued through 23 July with most of the data collected in late May through July. There were a total of 240 hours of data available. As described in the previous report (Scherrer and Wilcox, 1983) the data were scaled for each observing run to

have the same daily variance. The daily normalization factor used is highly correlated to the integrated signal power in the 5-minute band, suggesting that the variations in power are due to instrumental changes. Although the analysis was done using this daily normalized data, the results were rescaled using the average normalizing factor to provide an approximate scale in m/s. This normalization has both desired and undesired effects. It allows combining data from different years and from different observation sets where the raw data shows jumps in average signal due to variations in the instrument with time and to variations in the preliminary data reduction procedures. This method also removes apparent sensitivity changes that may result from variations in solar activity. The main structure of the spectrum is the same for both the normalized and raw data, so the normalized data were used to allow direct comparisons with different intervals of time.

### *Significance of Spectrum.*

An estimate of the number of peaks in the spectrum that are significantly above the noise can be made by examining the cumulative distribution of the spectrum. A power spectrum computed from a normally distributed noise source will have an exponential distribution, thus the logarithm of the cumulative distribution will decrease linearly with power. A departure from a straight line is an indication of the presence of significant spectral features and the slope of the line is a measure of the variance ( $\sigma^2$ ) in the spectrum.

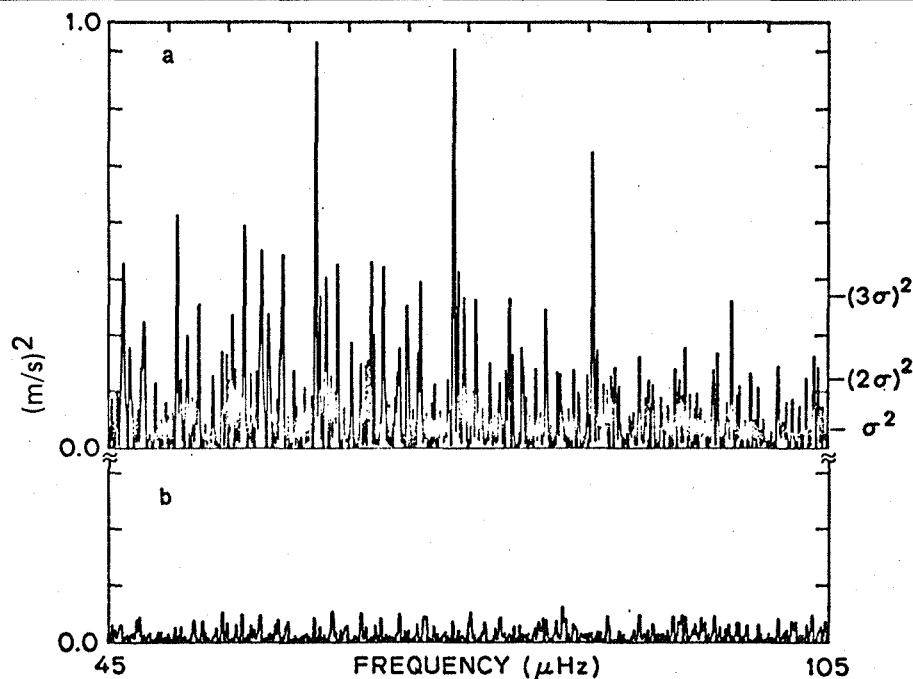


Figure 1. The power spectrum of velocity observations from the Stanford Solar Observatory in 1979. The spectrum in the range 45-105  $\mu\text{Hz}$  (360 to 160 minutes) is shown. Part (a) shows the original spectrum, part (b) shows the spectrum to the same scale after fourteen peaks were identified and the associated sinusoidal waves subtracted from the data. The scale shown has been corrected for the average normalizing factor. From Delache and Scherrer (1983).

Figure 2 shows this plot for the 45-105  $\mu\text{Hz}$  range of the 1979 spectrum. The actual data is represented by large dots. The smaller points are for a spectrum computed the same way from data constructed by taking the original data for each day in reversed order, thus keeping the original window function. Any coherent signal present in the original data will be eliminated by this procedure. The noise statistics for the modified data power spectrum should be unchanged for periods up to the average daily observing time, i.e. for frequencies higher than 40  $\mu\text{Hz}$ . It can be seen that in the actual data there are more peaks with value above  $0.25 (m/s)^2$  than in the reversed data. The level at which the actual distribution begins to depart from the noise distribution is around  $2.5 \sigma$ . Since the spectrum was computed with a resolution of 0.02  $\mu\text{Hz}$  but has a natural resolution of only 0.11  $\mu\text{Hz}$ , there will be about five points shown for each significant peak in the spectrum. Also, since the data has gaps at night, the day sidelobe structure introduces 2 to 4 apparently significant artificial peaks for each true peak. This reduces the number of independent spectral estimates to about 200. These considerations suggest that there are about 10 independent peaks above the  $2.5 \sigma$  level.

*Cleaning the Spectrum*

Since the spectrum appears to be dominated by the sidelobe structure from the observing times, a procedure must be performed to eliminate these sidelobes.

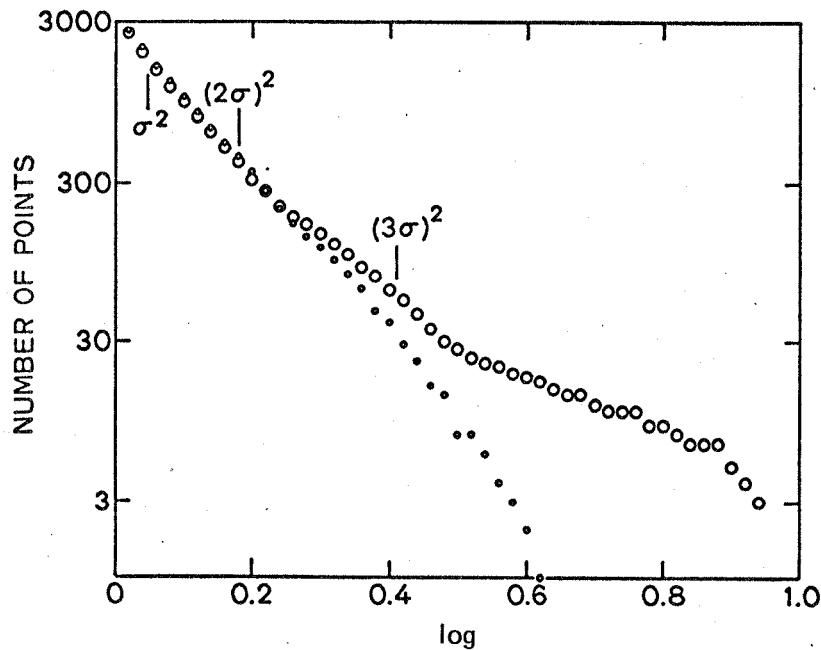
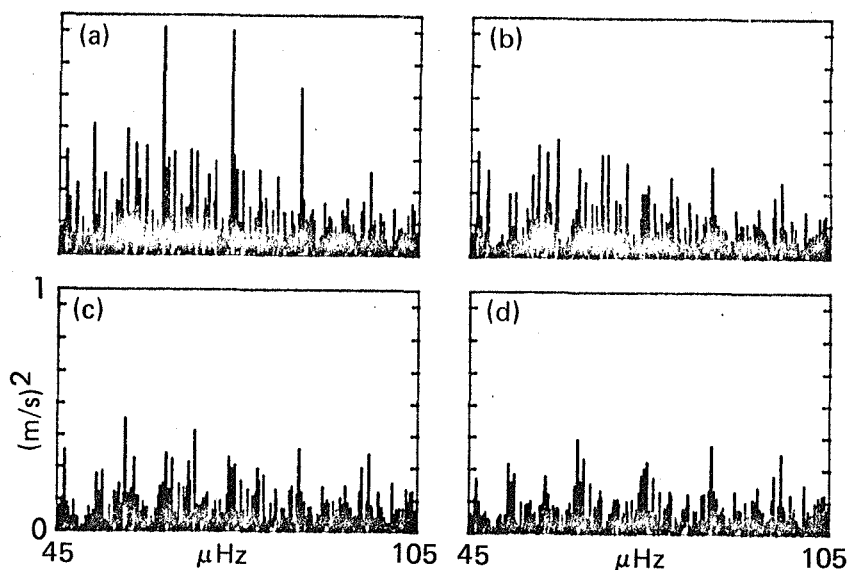


Figure 2. The cumulative distribution function of the spectrum in Figure 1a. The log of the number of spectral estimates with size larger than a given value is plotted vs that value. The large dots correspond to the observed spectrum. The small dots are from a similar spectrum computed with the data within each day taken in reversed order. The small dots then refer to a spectrum with the same window and noise characteristics but no coherent oscillations. The variance which is deduced from the slope is  $\sigma^2 = (20 \text{ cm/s})^2$ .

Figure 3. The spectrum of the original data (a), the spectrum after removing the largest peak (b), after removing the second peak (c), and after removing the third peak (d). (from Scherrer, 1983).



freq. $\mu\text{Hz}$	period min.	power $(\text{m/s})^2$	$l$	$n$
58.25	286.1	0.15	1	10
59.52	280.0	0.52	1	10
64.13	259.9	0.17	1	9
65.30	255.2	0.22	1	9
73.43	227.0	0.18	1	8
*73.84	225.7	0.98	1	8
83.50	199.6	0.09	1	7
95.70	174.1	0.23	1	6
96.94	171.9	0.31	1	6
*50.72	328.6	0.07	2	20
56.33	295.9	0.49	2	18
59.52	280.0	0.52	2	17
63.12	264.1	0.42	2	16
66.65	250.0	0.13	2	15
46.23	360.5	0.25	day/4	
62.29	267.6	0.98	-	
92.21	180.7	0.10	day/8	

Table 1. Frequencies of the 14 peaks found in the 45-105  $\mu\text{Hz}$  range of the 1979 spectrum. The frequencies are in  $\mu\text{Hz}$  with the corresponding periods shown in minutes. The power is shown as amplitude squared  $(\text{m/s})^2$ . The classification of  $n$  and  $l$  are described in the text. The 2 peaks marked with an asterisk are an alternate identification for the peak at 62.29  $\mu\text{Hz}$  and its associated side-lobes. Note that the peak at 59.52  $\mu\text{Hz}$  could be identified as part of either an  $l=1$  or  $l=2$  series and is included in the table twice.

An iterative peak removal technique was used to find and remove the peaks one at a time. First an FFT was computed and the largest peak determined. Next that peak was accurately found with a fine resolution simple Fourier transform in the vicinity of the peak. Finally the corresponding sinusoidal signal was subtracted from the original data. This procedure was repeated, producing a list of frequencies free of day sidelobes. Figure 3 shows the original spectrum and the spectrum after removing each of the first 3 peaks. It can be seen that for each peak removed, the entire sidelobe structure is also removed. Selecting peaks with amplitudes greater than about 30 cm/s, 14 peaks were found in the frequency band 45 to 105  $\mu\text{Hz}$ . Using the same amplitude criterion, another 8 peaks were found at frequencies in the range

extending to 165  $\mu\text{Hz}$ . These higher frequency peaks will not be discussed here. The list of 14 peaks is given in Table 1. Note that since the spectrum is assumed to consist of a set of line spectra, the scaling used in this paper is amplitude squared rather than power in the usual sense. To convert to the usual power spectrum scaling all values must be divided by 4 and would refer to intervals on 0.11  $\mu\text{Hz}$  (see e.g. Bath, 1974, p183).

To see the effect of removing the complete set of peaks from the original data, the spectrum of the residual data after subtracting these 22 sinusoids is shown in Figure 1b. Figure 1b is shown to the same scale as Figure 1a. It is clear that far more than 14 peaks in the band shown have been removed in the resulting spectrum. This method of peak identification appears to be a powerful tool for analysis of power spectra computed from data with complicated windows. The method finds a table of periods, phases and amplitudes that best clean the spectrum. From experiments with artificial data, we have concluded that while the correct peak is usually found, a side lobe introduced by the window is occasionally selected. This is a weakness of the procedure that requires care in examining the resulting peaks. We have also found that *the amplitudes determined from data with a complicated window are not to be relied upon.*

The method described above relies upon the single assumption that the examined signals consist of coherent sine waves. It is then not surprising that for each frequency removed, a complete day-sidelobes set of lines disappears from the spectrum. A further check is necessary to determine whether or not the signal is coherent for the full span of the observations. For that purpose, the data with the tabulated frequencies removed was divided into two parts, 7 April through 15 June and 16 June through 23 July. Figure 4 shows a set of nine spectra resulting from this division. The top row shows the original spectrum (left), the spectrum of the first half of the data (middle), and the spectrum of the last half of the data (right). A large variation in the appearance of the spectra is obvious. A first impression is that these spectra represent either very noisy data or unrelated data. Examination of the bottom two rows shows that the first impression is wrong. The second row contains the residual spectra similar to Figure 1b for each of the intervals in the top row. The middle row, left spectrum is essentially the same as Figure 1b. The center and right frames are for two halves of the interval but cleaned by subtracting the sinusoids found for the *entire* interval. The spectra of each part separately do not show any of the removed frequencies. If any of the frequencies removed were not present in both halves of the data, they would have been artificially put into the cleaned data by the subtraction of the sinusoid over the entire interval. In that case they would show in the spectrum of one or both halves of the cleaned data. This is strong evidence that the signals are all present in both halves of the observations. The bottom row demonstrates the variability in amplitude caused by the observing window. Using only the times of observations and the list of frequencies and phases found above, a cleaned dataset was reconstructed. The spectra from the entire interval and the two halves are shown in the bottom row. Note that all three frames are computed from artificial data generated from the same set of frequencies and phases. The only difference between frames on the bottom row is the observing window. The conclusion is that for windows obtained at Stanford (which are similar to the average distribution of observations in the Crimea and at SCLERA), the relative amplitudes of peaks may be dominated by interference caused by data gaps.

#### *Peak Identification*

Returning to table 1, two of the 14 peaks are apparently day aliases at 1/4 and 1/8 of a day. (With only 3 months of data the previously reported peak at 160.01 minutes can not be distinguished from 1/9 of a day and is masked by the sidelobes

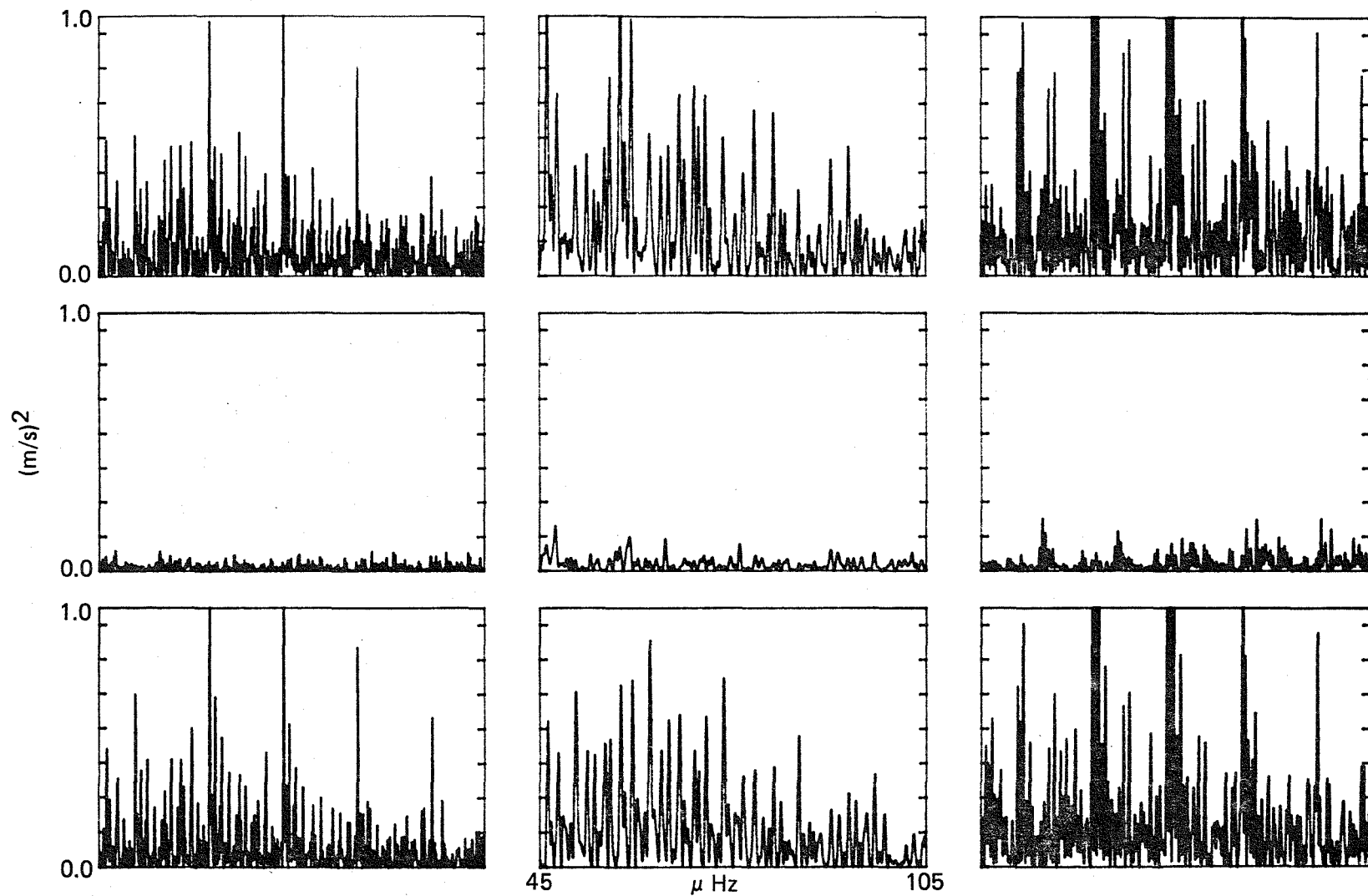


Figure 4. Spectra of the entire interval and both halves of the interval. The first column contains spectra for the entire data interval. The center column is for the first half, the right column is for the last half. The top row is the original data, the middle row is the residuals after subtracting the peaks found in the entire interval (similar to Figure 1b), and the bottom row is data reconstructed from the peak list.



of the lower day-harmonics). This leaves 12 peaks for further analysis. There is some ambiguity in the true identification of the largest of the remaining peaks. The first 1/day sidelobe is of the same amplitude as the main peak. This could be due to two true peaks with separation 1/day or 2/day. By careful examination of the peak shapes, the peak at  $62.29 \mu\text{Hz}$  has been identified as the true peak with a smaller peak at  $96.94 \mu\text{Hz}$ , although the identification could have been made as a true peak at  $73.84 \mu\text{Hz}$  with smaller peaks at  $50.72$  and  $96.94 \mu\text{Hz}$ . The alternate identifications are shown in parentheses in Table 1.

### Interpretation

In order to interpret the peaks found in this spectrum, one must seek guidance from theory. Estimates for the spacing of  $g$ -mode oscillations made from the full asymptotic approximation (Tassoul 1980) and calculated from complete solar models (Berthomieu, Provost, and Christensen-Dalsgaard, private communications) suggest that the  $g$ -mode oscillations of the same degree  $l$  should be about equally spaced in period for order large enough ( $n > 6$ ). Therefore the list of prominent peaks was examined for equal spacing in period. Three of the four largest peaks are separated in period by 15.5 minutes. This is very close to the spacing of the standard solar model for  $g$ -modes with degree two and order greater than 10.

To aid in further identification of the modes the asymptotic approximation (Tassoul 1980) is used. This approximation shows that for sufficiently large order  $n$  the period  $T$  is:

$$T = T_0 \left( n + \frac{l}{2} - \frac{1}{4} \right) / \sqrt{l(l+1)}$$

If the three peaks with 15.5 minute separation are part of an  $l=2$  series, then  $T_0$  and

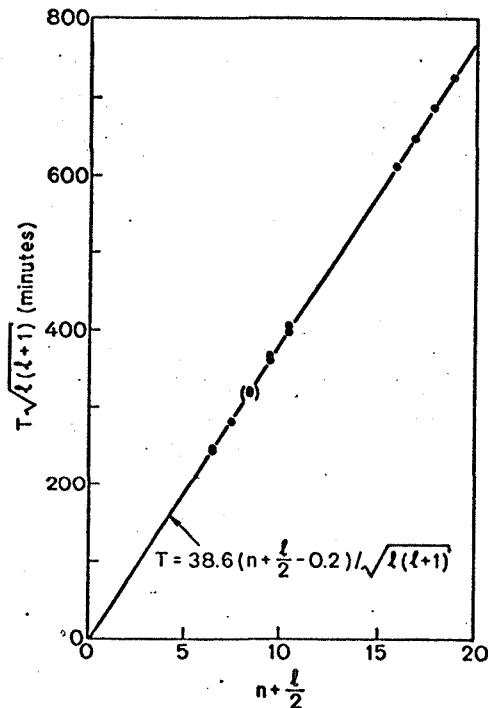


Figure 5. The observed modes are plotted within the structure of the asymptotic formula described in the text. The line is determined from the three largest peaks that have been assigned. From Delache and Scherrer (1983).

thus the probable spacing for the  $l=1$  and  $l=3$  series can be computed. Doing this, it is seen that most of the remaining peaks are likely to be part of the  $l=1$  series. Using the estimate for  $l$ , the observed periods can be plotted as  $T\sqrt{l(l+1)}$  vs  $(k + l/2 - 1/4)$  where  $k$  is an integer increasing with period. If the peaks are consistent with the model and the correct degree  $l$  has been assigned, the points will all lie on a straight line. From the intercept of this line both the best value for  $n$  and the correct value for  $-1/4$  term can be determined. This has been done in Figure 5.

The three largest peaks were used to define the  $l=2$  series and thus to find  $T_0$ . The rightmost dots in Figure 5 represent the three largest peaks which determine the  $l=2$  series. The other modes identified in Table 1 are also shown. The line was found from the three large  $l=2$  peaks only. It can be seen that the other peaks are organized in the expected  $g$ -mode structure. The  $T_0$  implied by the identified modes is  $38.6 \pm 0.5$  minutes. The intercept of the best fit line with the abscissa is expected to be 0.25 and is found to be  $0.2 \pm 0.1$ , thus the order  $n$  is most likely correctly determined. If  $n$  is decreased by one for each peak,  $T_0$  becomes about 41 minutes but the scatter is increased somewhat.

The largest peak at  $62.29 \mu\text{Hz}$  appears not to be consistent with the asymptotic formula with the  $T_0$  we have found. If the alternate choice of the  $62.29 \mu\text{Hz}$ - $73.84 \mu\text{Hz}$  pair were chosen, the largest peak would be at  $73.84 \mu\text{Hz}$  which would be identified as  $l=1$ ,  $n=8$ . In this case a peak at  $50.7 \mu\text{Hz}$  shows up in the peak finding procedure and would be identified as  $l=2$ ,  $n=20$ .

Note that in several cases two peaks have been assigned to one order  $n$  in the  $l=1$  series. These peaks have an average separation of  $1.2 \mu\text{Hz}$  and could be evidence of rotationally split modes. Assuming the  $l=1$  modes are  $m=\pm 1$ , the  $1.2 \mu\text{Hz}$  separation would correspond to a rotation rate  $\Omega=7.5 \times 10^{-6} \text{sec}^{-1}$ . The rotational splitting kernel for these modes has a broad maximum between 0.2 and 0.5  $R$ . Assuming a smooth form for  $\Omega(r)$  as in Figure 2 of Gough (1982),  $J_2$  can be calculated and is of order  $1.7 \times 10^{-6}$ . This value differs from that found by Hill et al. (1982).

The sensitivity of the Stanford instrument to degree  $l=2$   $g$ -mode oscillations must be considered. For the previously reported acoustic mode oscillations in the 5-minute range, the Stanford velocity differencing scheme is most sensitive to modes of degree  $l=3$  to  $l=5$  but is not sensitive to modes of degree  $l=1$  or  $l=2$ . For low frequency  $g$ -modes, Gough and Cristensen-Dalsgaard (1982) have shown that for periods around 160 minutes the Stanford instrument is most sensitive to modes with degree  $l=5$  to  $l=7$ . For longer periods however, the instrument begins to be relatively more sensitive to modes with lower degree. This variation with period, unlike the case for the  $p$ -modes, is due to the importance of the horizontal motions induced by the internal gravity modes. To properly estimate the sensitivity to the various modes the oscillation energy must also be considered. In the simplest case, the combined effect of equipartition of energy between modes and the mode visibility with the Stanford instrument yields the greatest sensitivity for  $l=1$  and  $l=2$  at these longer periods.

The theory predicts several hundred periods for low degree  $g$ -modes in the period interval examined. Only the most visible handful of peaks has been examined and has been found to be consistent with the theory. The relative peak sizes and simplicity of the analysis leads to the present mode identification. Alternative identifications are quite possible but they can only be tested with more data from other years or from other observatories.

### *Note Concerning Lifetimes*

There are some general considerations concerning the appearance of power spectra of data containing oscillations of finite lifetime that must be remembered when interpreting such spectra. It is generally recognized that the spectrum of an oscillator with a lifetime  $\tau$  greater than the observation time  $T$  will be a "line" that will appear as a  $\text{sinc}^2$  with width  $1/T$ . It is also recognized that an observation of a single oscillation event with lifetime  $\tau$  and observation time  $T$  will result in a similar spectral line but with width  $1/\tau$ . What is sometimes forgotten is that if the observation duration includes several, but not many instances of an oscillation lifetime, i.e.  $T > \tau$  but not  $T \gg \tau$ , the spectrum will be a complex structure of lines of width  $1/T$  in an envelope of width  $1/\tau$ . If  $T = 2\tau$  the spectrum will look similar to that of a rotationally split  $l = 1$  mode. Figure 6 shows spectra computed from artificial data constructed with  $T = 3$ -days through  $T = 100$ -days and  $\tau = 3$  days through  $\tau \gg 100$  days. There were no data gaps assumed. The plot scales were adjusted such that the plots almost fill the frame. An examination of Figure 6 shows that caution must be used in interpreting spectra of oscillatory signals of unknown lifetimes.

### *Acknowledgements.*

I thank Philippe Delache for acting as an ambassador of  $g$ -modes. His efforts have resulted in timely comparisons of results from the available sets of observations. This work was supported in part by the Office of Naval Research under Contract N00014-76-C-0207, by the National Aeronautics and Space Administration under Grant NGR05-020-559 and Contract NAS5-24420, by the Atmospheric Sciences Section of the National Science Foundation under Grant ATM77-20580 and by the Max C. Fleischmann Foundation.

### *References.*

- Bath, Markus, *Spectral Analysis in Geophysics*, Elsevier Scientific Publ. Co., (1974).  
Berthomieu, G., J. Provost, and J. Christensen-Dalsgaard, (private communications).  
Christensen-Dalsgaard, J. and D.O. Gough, *Mon. Not. Roy. Astron. Soc.*, **198**, 141 (1982).  
Delache, P. and P.H. Scherrer, *Nature*, **306**, 651.  
Gough, D. O., *Nature*, **298**, 334, (1982).  
Hill, H. A., R. Bos, and P. R. Goode, *Phys. Rev. Lett.*, **49**, 1794, (1982).  
Scherrer, P.H., *Proceedings of the Study Conference "Oscillations as a Probe of the Sun's Interior"*, Catania, (submitted 1983).  
Scherrer, P.H., Wilcox, J.M., Kotov, V.A., Severny, A.B. and Tsap, T.T., *Nature*, **277**, 635-637 (1979).  
Scherrer, P.H., J.M. Wilcox, J. Christensen-Dalsgaard and D.O. Gough, *Solar Phys.* **82**, 75-87, (1983).  
Scherrer, P.H. and Wilcox, J.M., *Solar Phys.*, **82**, 37-42 (1983).  
Tassoul, M., *Astrophys. J. Suppl.*, **43**, 469-490 (1980).

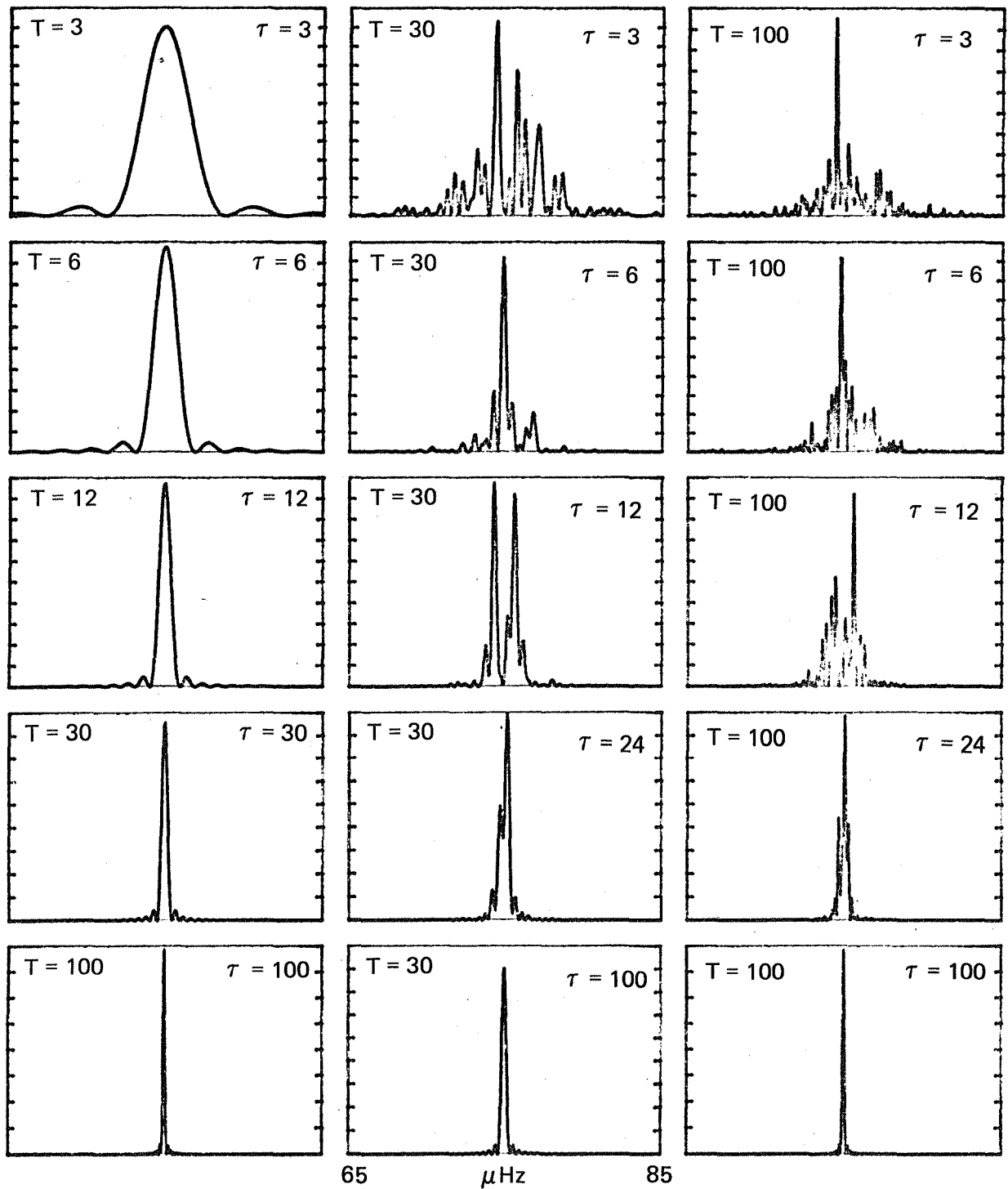


Figure 6. Artificial spectra computed from a single frequency signal with a selection of sample times and lifetimes. The sample duration  $T$  and mode lifetime  $\tau$  vary from 3 days to 100 days.

## SOLAR GRAVITY MODES FROM ACRIM/SMM IRRADIANCE DATA

Claus Fröhlich<sup>1</sup> and Philippe Delache<sup>2</sup>

<sup>1</sup>Physikalisch-Meteorologisches Observatorium, World Radiation Center, Davos, Switzerland;

<sup>2</sup>Observatoire de Nice, Nice, France

### ABSTRACT

The record of 280 days of continuous data of the ACRIM radiometer on board the Solar Maximum Mission satellite is analysed in the frequency range from 10 to 80  $\mu\text{Hz}$ . Gravity modes of degree one and two with orders from about 10 to several hundreds can be localized.

A statistical method to determine the fundamental period  $T_0$  and the rate of rotation  $\nu_R$  as seen by rotational splitting is described and the results for  $33.5 < T_0 < 45.5$  minutes and  $0.4 < \nu_R < 2.0$   $\mu\text{Hz}$  presented. They indicate a rather high  $T_0$  and it cannot be excluded, that it is above the upper limit analysed.

### INTRODUCTION

After the detection and identification of solar gravity modes by Scherrer and Delache (1983) it became obvious, that the ACRIM-SMM record of 280 days during 1980 would be an ideal test for the identification, if such modes could be seen as irradiance fluctuations. Furthermore, it would allow us to extend the accessible range of frequencies substantially to lower frequencies as no daily observational window deteriorates the power spectrum below about 30  $\mu\text{Hz}$ .

The theory of gravity modes predicts a visibility in irradiance at frequencies below about 60  $\mu\text{Hz}$  (Berthomieu, 1983). Thus, the time series of orbital means of the SMM data is just ideal for this kind of analysis, as the Nyquist frequency of this sampling is 86  $\mu\text{Hz}$ . A first attempt to search for g-modes in the SMM/ACRIM data has been presented by Fröhlich and Delache (1984, in the following referred as F&D). In the present paper the method is refined and the range of analysis substantially extended.

## SMM-ACRIM DATA SET AND POWER SPECTRUM

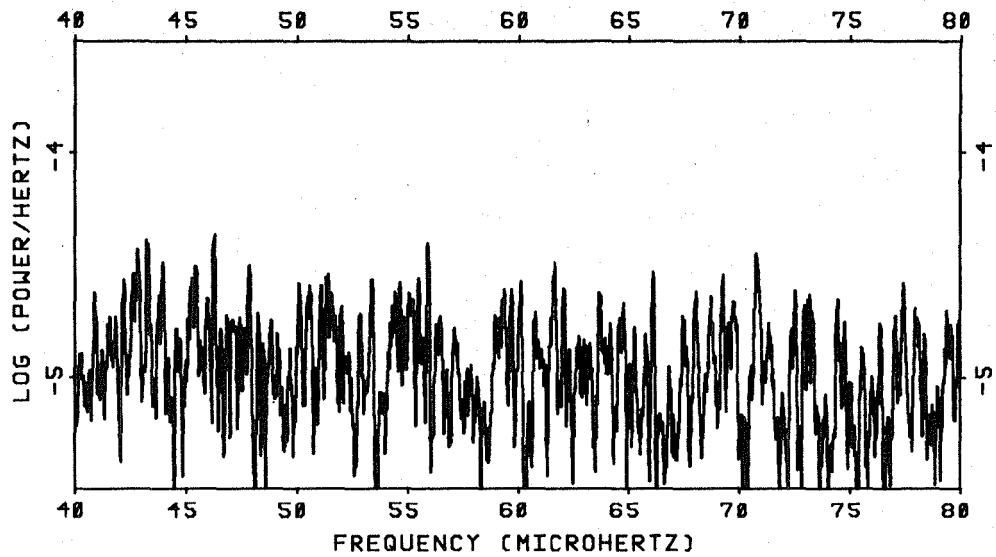
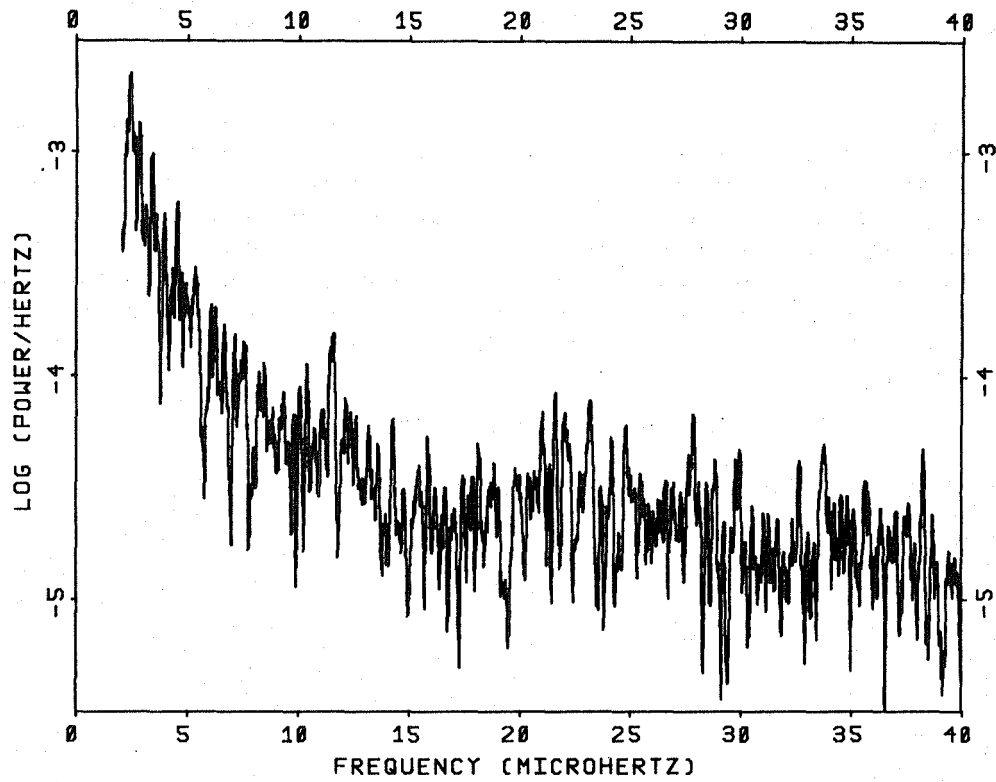
The excellent data set of the ACRIM radiometer consists of individual readings of the irradiance with a sampling of roughly two minutes. Due to the orbit of SMM around the earth, data are lacking for about 30 minutes out of 96 minutes. In order to get a homogeneous data set a strictly periodic triangular weighting function with a length of the base of the triangle of 66 minutes and a period of 96 minutes was used to calculate "orbital means". Further details on the data themselves and the methods used to ensure the quality may be found in F&D.

For the frequency analysis the FFT technique is used and the data are supplemented by zeros in order to oversample the frequency by a factor of about four. This is to allow an unbiased estimate of the location of the peaks in the spectrum. The natural resolution of the spectrum is  $0.04 \mu\text{Hz}$ . Figure 1 shows a smoothed spectrum for the range from 2 to  $80 \mu\text{Hz}$ . An obvious feature of the spectrum is its increase toward lower frequencies. From the highest frequencies down to about  $15 \mu\text{Hz}$  it follows roughly a  $1/f$  behaviour and turns toward lower frequencies to a  $1/f^2$  law. The latter may be due to g-modes as their number density per unit frequency interval also increases with  $1/f^2$  and the density of the modes above about  $15 \mu\text{Hz}$  is already so high that they can no longer be resolved. In order to assess the statistical significance of the peaks, in F&D a power density analysis of the unsmoothed spectrum was performed. The results for the frequency range  $10\text{--}80 \mu\text{Hz}$  show a mean deviation from the straight line at a power of about  $50 \cdot 10^{-6} \text{ Hz}^{-1}$ . This result means, that about 120 peaks of the unsmoothed spectrum extend beyond this limit and are thus significant in the sense that they are larger than the ones one would expect from a pure noise spectrum.

## STATISTICAL METHOD FOR G-MODE SEARCHING

At frequencies where the g-modes are expected to be seen in the irradiance their number density for each degree is already high and the structure significantly complicated by the rotational splitting, the limited resolution and the noise of the spectrum. Therefore, it would be a very tedious and complicated task to try to identify each peak individually and it seems that only a statistical method can be used successfully.

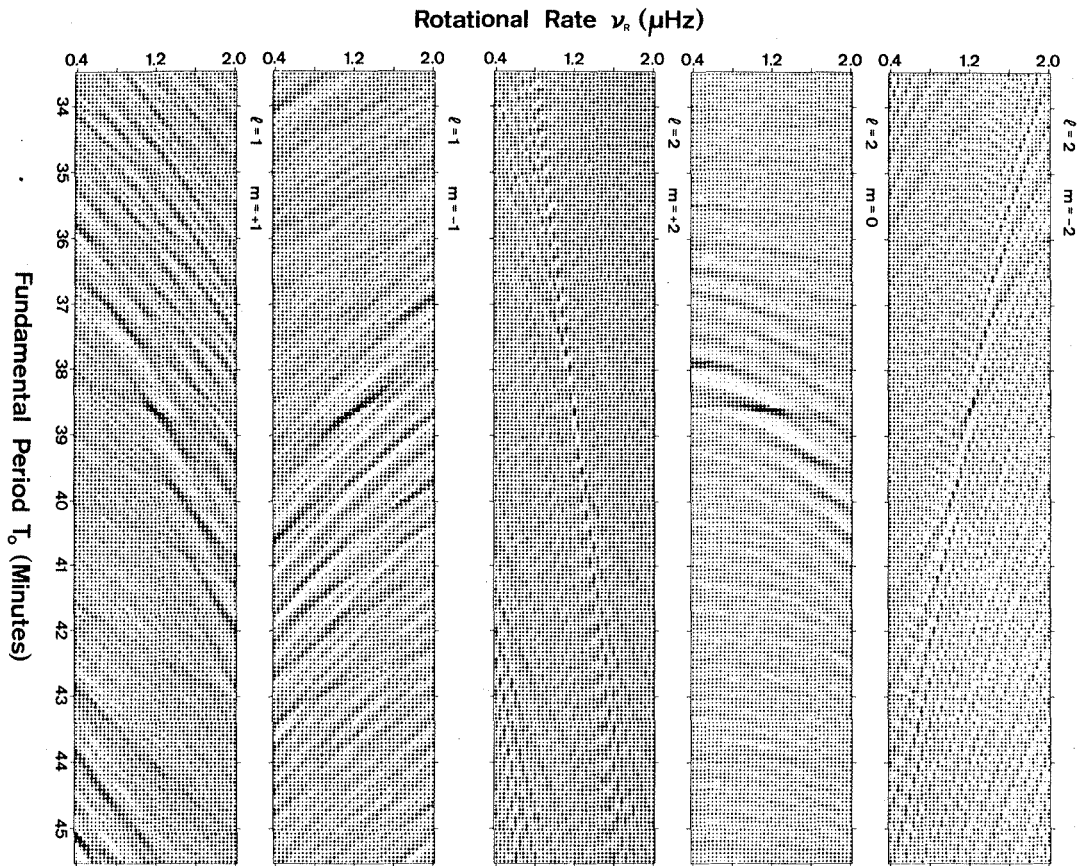
The order of the g-modes in the frequency range below  $80 \mu\text{Hz}$  is always higher than about 10 to 20 for  $l=1$  and 2. Therefore, an asymptotic behaviour as described by Tassoul (1980) could be assumed. This was the basis for the analysis by F&D. However, as the periods considered can no longer be regarded as short compared with the rotational period itself, coupling between the modes has to be taken into account. Berthomieu et al (1978) have developed the necessary formalism to treat this effect. Results show that at the limit of low radial orders the effect is negligible and the Tassoul formula is still valid. For periods above about  $40 \mu\text{Hz}$  the deviations become important. Thus all frequencies of the g-modes are now determined by directly calculating the eigenvalues of the trun-



**Figure 1:** Smoothed power spectrum of 280 days of ACRIM irradiance measurements in the range from 2 to 80  $\mu\text{Hz}$ . Smoothing reduces the natural resolution to 0.1  $\mu\text{Hz}$ .

cated infinite tridiagonal matrix of Berthomieu. To reduce the computer time a two dimensional interpolation scheme with a  $4 \times 5$  matrix of function values covering the envisaged range of  $T_0$  and  $\nu_R$  for each  $\ell, m, n$  combination is used in the program of the analysis.

For a fixed fundamental period  $T_0$  and a fixed rotational rate  $\nu_R$  the line positions of the g-modes of different orders  $n$  for each  $\ell, m$  combination are calculated and bins are filled with the power from the spectrum to be analysed within the width of the natural resolution and centered at this mode frequency. The powers of all orders  $n$  of one mode of one  $\ell, m$  combination within the frequency range concerned are summed up and a mean power per mode  $\ell, m$  is calculated. By varying  $T_0$  and  $\nu_R$  five two dimensional arrays of mean power per mode can be constructed ( $\ell=1, m=-1, +1; \ell=2, m=-2, 0, +2$ ).



**Figure 2:** Analysis of a synthetic spectrum with  $T_0 = 38.6$  minutes and  $\nu_R = 1.2 \mu\text{Hz}$  for the modes  $\ell=1, m=\pm 1$  and  $\ell=2, m=\pm 2, 0$ . Each point corresponds to the mean power per mode at the corresponding  $T_0$  and  $\nu_R$  and the darker the point the higher the power is.

As an example the results of the analysis of a synthetic spectrum of g-modes with  $T_0 = 38.6$  minutes and  $1.2 \mu\text{Hz}$  is shown in Figure 2. The synthetic spectrum is noise free and a constant power of 1 is attributed



to each mode  $\ell, m, n$  at the frequency determined with the method described above. As expected, the two-dimensional pictures show one peak for each mode  $\ell, m$  at 38.6 minutes and 1.2  $\mu\text{Hz}$ . But they show also some structure outside the peaks. This is partly due to the fact, that in a limited frequency range a change in  $T_0$  can be compensated by a change in  $\nu_R$  and vice versa and partly due to some interference of modes not corresponding to the mode analysed. An obvious example of the first kind is the slant ridges in the  $\ell=1$  diagrams and the curved ridges in the  $\ell=2, m=0$  ones. The curvature of the latter is due to the coupling between the modes; the asymptotic approximation would yield a rotation independent  $\ell=2, m=0$  mode.

The same analysis has been performed for a simulated noise spectrum with no g-modes incorporated. For the simulation a random amplitude and phase of the signal are assumed. This yields a  $\chi^2$ -distributed power spectrum with a degree of freedom of two. Furthermore the mean power is assumed to vary as  $1/f$  in order to simulate the behavior of the ACRIM spectrum in the range between 10 and 80  $\mu\text{Hz}$  (cf Figure 1). The results are shown in Figure 3. Again a structure similar to one found in the synthetic data is displayed; however peaks in one diagram do no longer

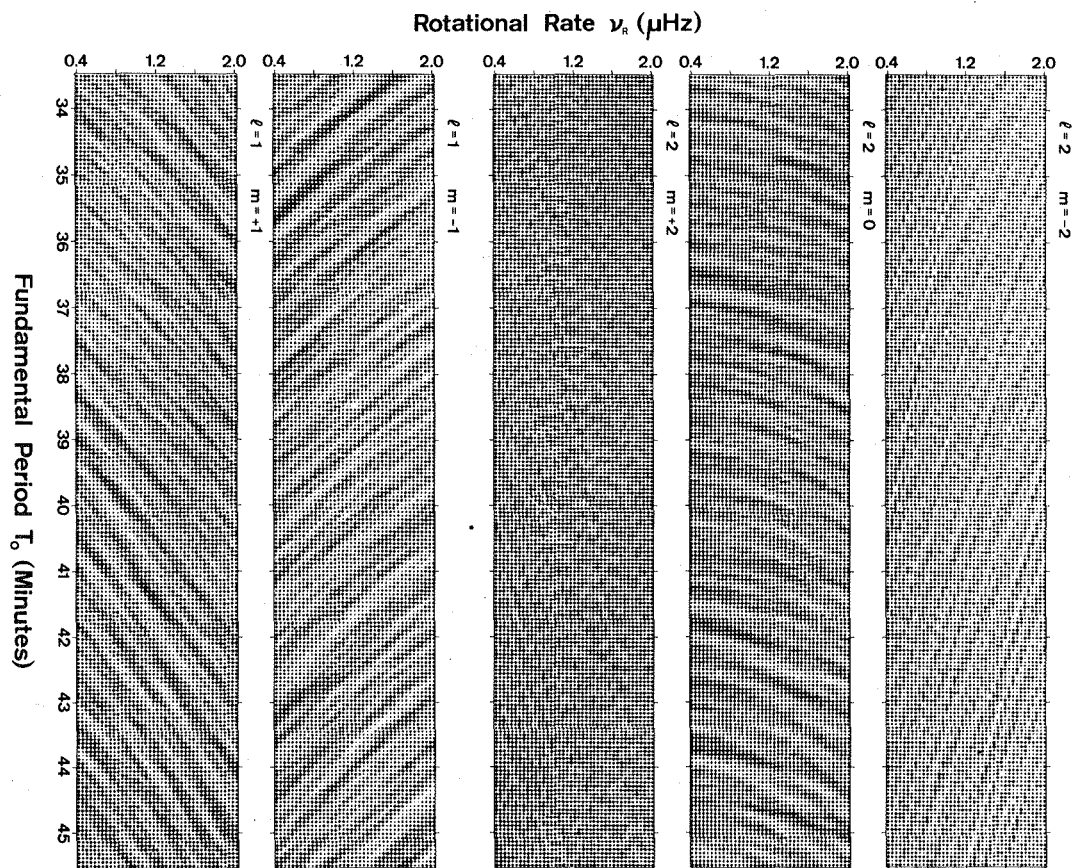
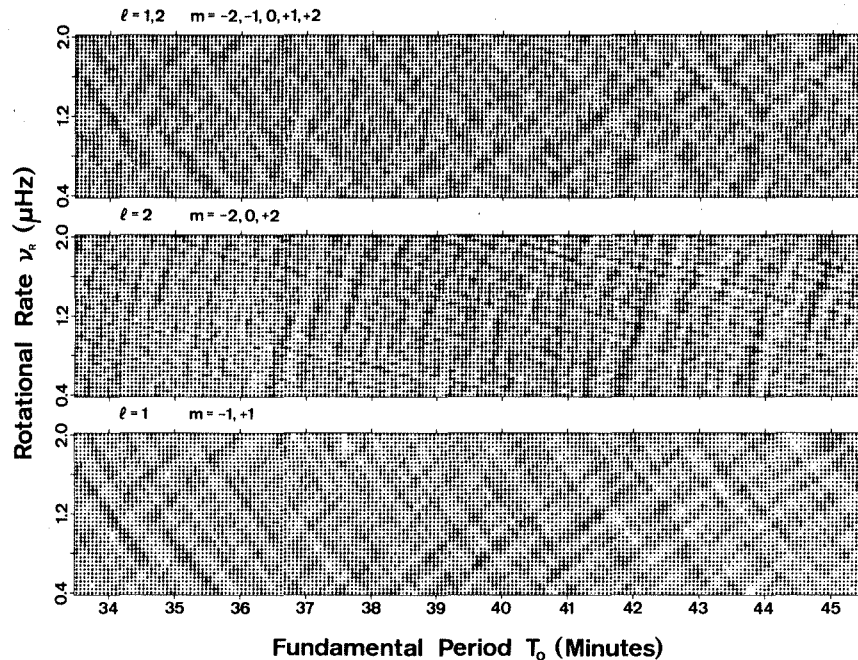


Figure 3: Same as Figure 2, but for a simulated noise spectrum in the range from 10-40  $\mu\text{Hz}$ .

correspond to peaks in the others. Moreover the structure in the rotationally split  $\ell=2$  diagrams is more or less lost. In order to reduce random peaks and to amplify the possible coincidence of peaks at a given  $T_0$  and  $\nu_R$  the different diagrams may be summed together. Figure 4 shows the corresponding results for the noise spectrum and indeed most of the individual peaks are now buried in the noise.



**Figure 4:** Diagrams of power, summed over the two  $\ell=1$ , the three  $\ell=2$  and all five together for a simulated noise spectrum in the range 10–40  $\mu\text{Hz}$ .

Similar diagrams can be calculated for the frequency range 40 to 80  $\mu\text{Hz}$ . As the number of modes in this range is much smaller, the peaks for the synthetic spectrum are now sharper and some blank space with no power at all is found in the diagrams. The noise spectrum on the other hand has less structure. Thus, this range would be much better suited for the determination of  $T_0$  and  $\nu_R$  than the range with the longer periods. But in contrast to the synthetic spectrum the amplitudes of the real spectrum decrease toward the higher frequencies.

## RESULTS AND COMPARISON WITH OTHER DETERMINATIONS

The results of the analysis of the ACRIM spectrum in the ranges 10–40 and 40–80  $\mu\text{Hz}$  are shown in Figures 5 and 7 for the individual modes and in Figures 6 and 8 for sums. The mean span, that is the mean difference between the minimum and maximum of a diagram in units of the mean power of the corresponding spectrum for the ACRIM is about 1.45 and 1.65 times greater than for the corresponding noise spectrum. The two numbers refer to the lower and higher frequency ranges. Thus, the ACRIM spectrum

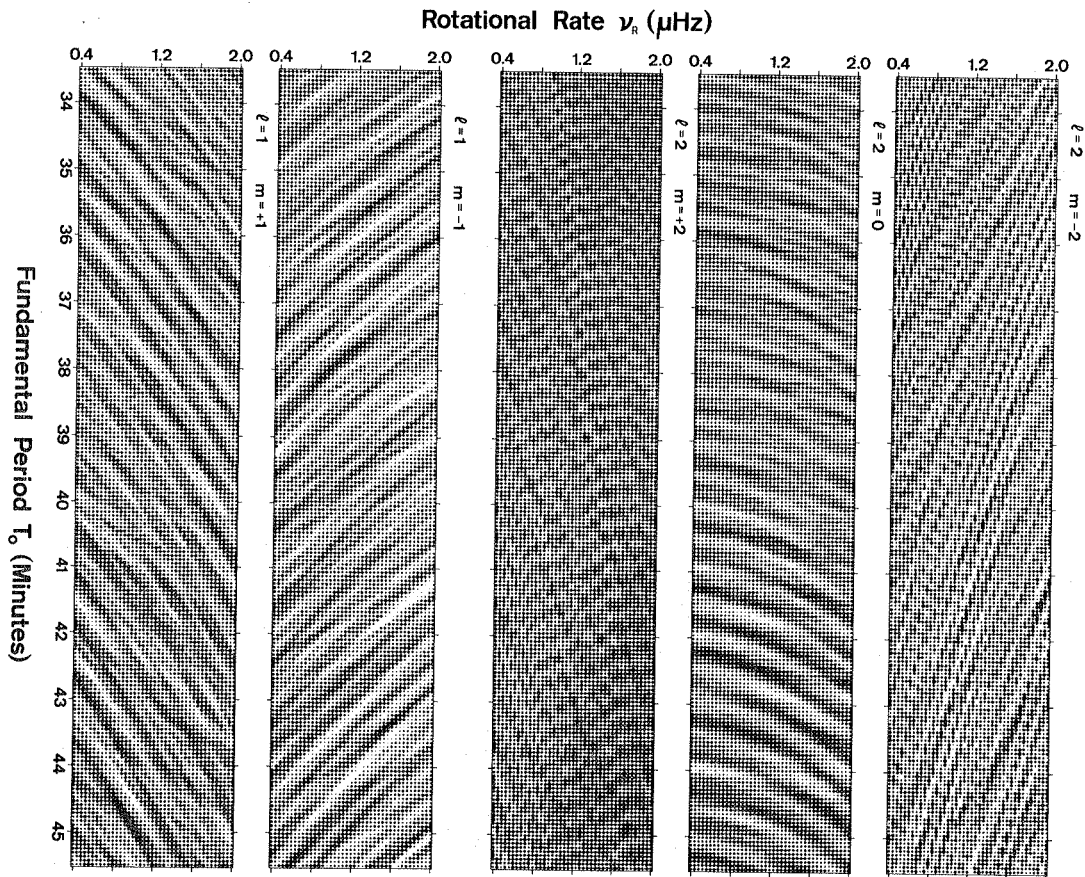


Figure 5: Same as Figure 2, but for ACRIM spectrum 10-40  $\mu\text{Hz}$ .

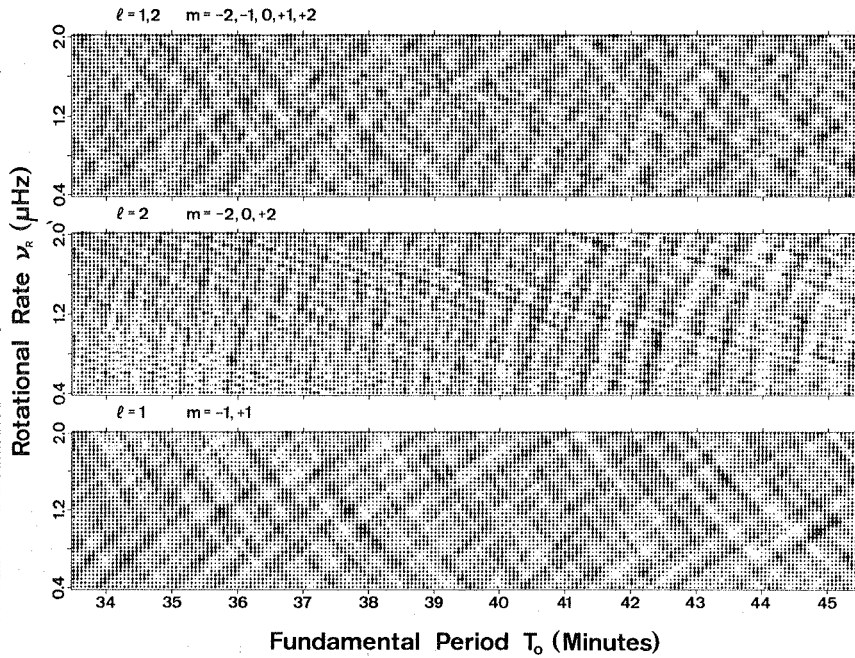


Figure 6: Same as Figure 4, but for ACRIM spectrum, 10-40  $\mu\text{Hz}$ .

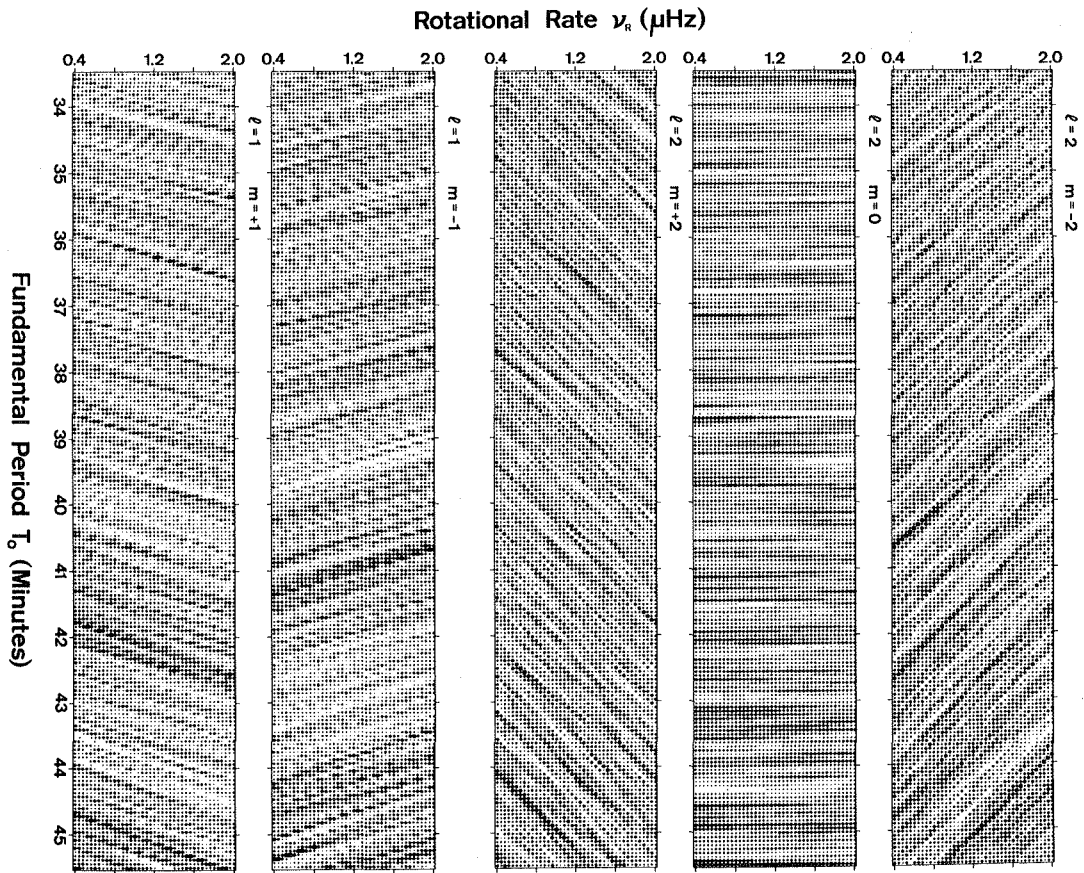


Figure 7: Same as Figure 2, but for ACRIM spectrum, 40-80  $\mu\text{Hz}$ .

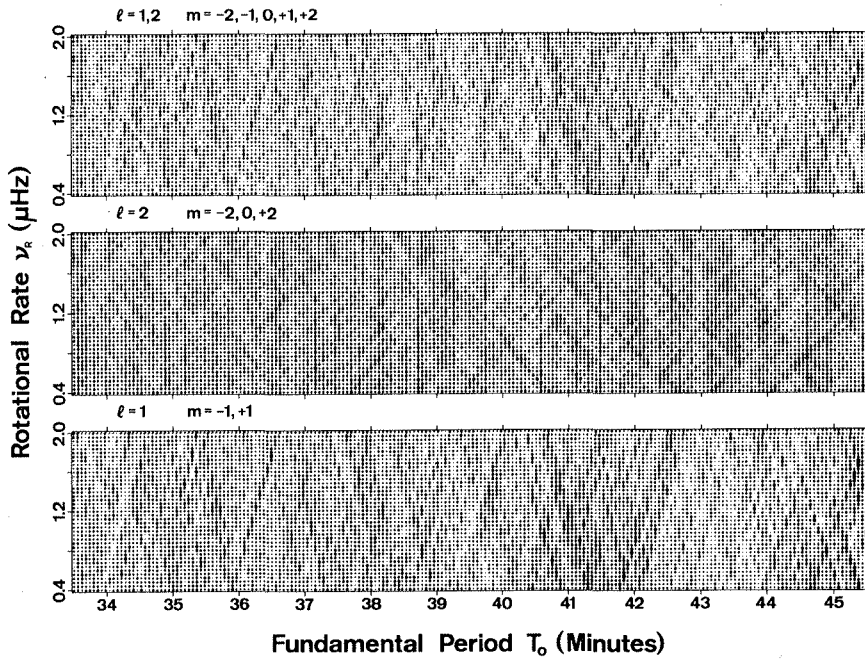


Figure 8: Same as Figure 4, but for ACRIM spectrum, 40-80  $\mu\text{Hz}$ .

shows significantly more structure than the one of simulated noise.

Several maxima are present in the diagrams. However, it is not obvious that one can be found, which occurs in all modes and the two frequency ranges at the same place. For the lower range the sum of the  $\ell=1$  diagram shows maxima at  $T_0$  of 35.7, 37.7, 38.1, 43.4 and 44.9 minutes. In the higher range only the one around 45 minutes is strong enough and found in both  $m$  values at the same place. This peak coincides with the one in the lower range and both are at a rotational rate of about 1.0  $\mu\text{Hz}$ . These results would therefore indicate a  $T_0 = 45.0 \pm 0.2$  minutes and a  $\nu_R = 1.00 \pm 0.05$   $\mu\text{Hz}$ .

As to the  $\ell=2$  the situation is less comfortable: In the lower range strong peaks are found around 42 minutes and somewhat weaker ones at 35.8, 41.5, 43.3 and 44.6 minutes. The diagrams of the higher range show a strong peak line just at the edge of the diagram at 45.5 minutes. Other smaller peaks may be localized at 37.2 and 38.7 minutes. But these are only present in the  $m=0$  and not in the  $m=\pm 2$  diagrams. In any case it seems, that at least for this frequency range the  $m=\pm 2$  modes are quite weak, if they exist at all.

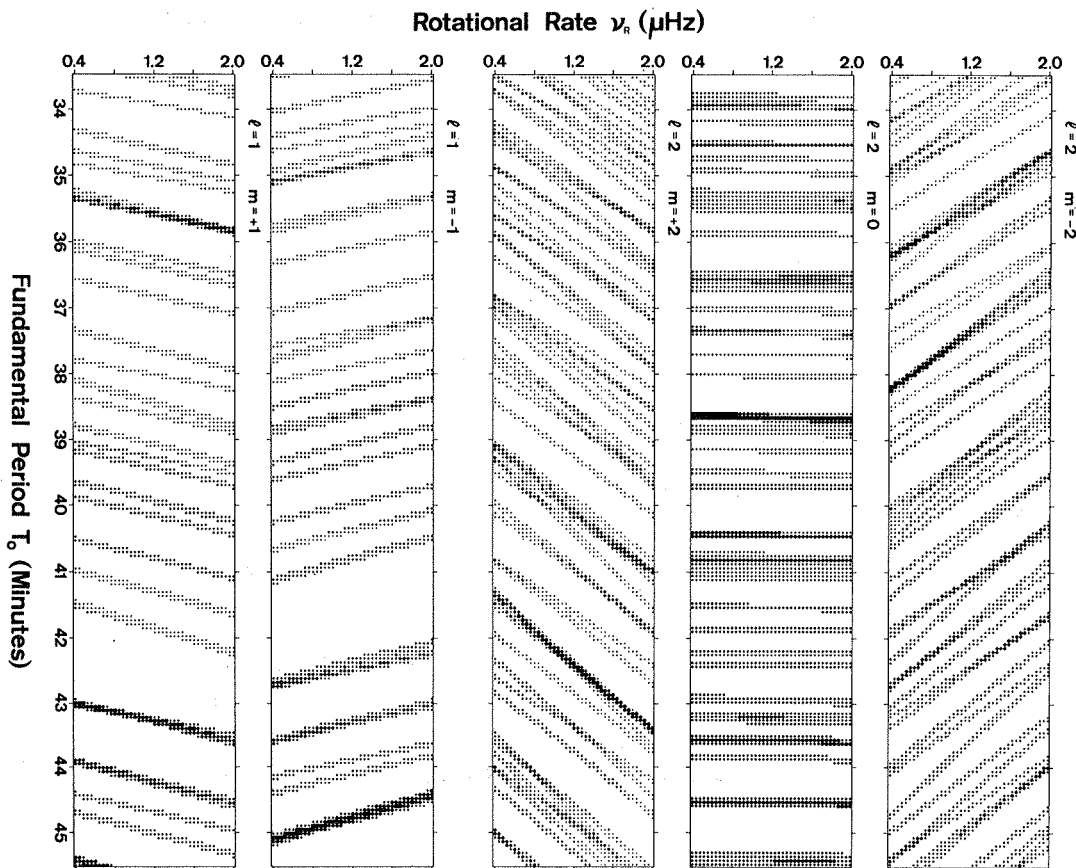


Figure 9: Same as Figure 2, but for the Stanford data, 40–80  $\mu\text{Hz}$ .

The only other determination of g-modes by Scherrer and Delache (1983) yields a  $T_0$  of 38.6 and  $\nu_R$  of about  $1.2 \mu\text{Hz}$ , which has been confirmed by the analysis of F&D in a very restricted range of  $T_0$  and  $\nu_R$ . This latter is obviously overridden by the present more extended analysis. In order to compare the Stanford with the ACRIM data a "synthetic" spectrum with amplitudes of 1 at frequencies of the lines below  $80 \mu\text{Hz}$  extracted from the Stanford spectrum is produced. The same type of analysis is made, but the bin width is increased to  $0.16 \mu\text{Hz}$  in order to account for the different resolution of the original data. The results are shown in the Figures 9 and 10. It is somewhat disturbing that only the  $\ell=2, m=0$  diagram shows a peak at 38.6 minutes. This may be due to the fact, that some lines - mainly identified as  $\ell=1$  - have been omitted, because they lie above  $80 \mu\text{Hz}$ . However, the resemblance with the ACRIM diagrams of the higher range is striking: For both the  $\ell=1$  peak lies towards the higher end of the diagram and no maximum is seen close to 38.6 minutes. On the other hand, both show the 38.6 peak line in the  $\ell=2, m=0$  diagram. These coincidences support the realism of most parts of the patterns. However, the nonexistence of a coincidence of the maxima in  $\ell=1$  and 2 is disturbing and may mean that the true maximum is outside the analysed range and probably above 45 minutes.

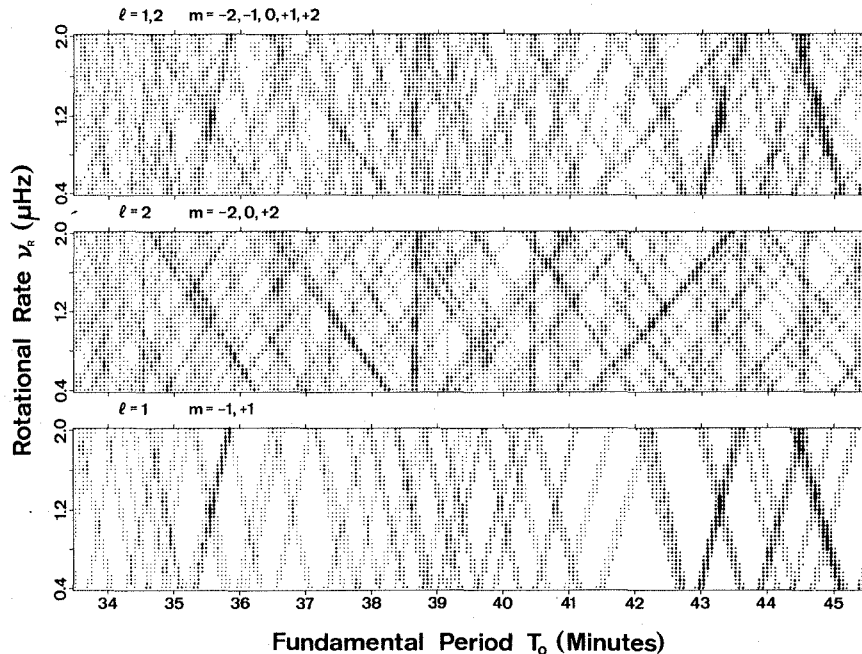


Figure 10: Same as Figure 4, but for the Stanford data, 40-80  $\mu\text{Hz}$ .

## REFERENCES

- Berthomieu,G., Gonczi,G., Graff,Ph., Provost,J., Rocca,A.: 1978, *Astron. Astrophys.* **70**, 579.
- Berthomieu,G.: 1983, Observatoire de Nice, *private communication*.
- Fröhlich,C., Delache,P.: 1984, *Memorie della Società Astronomica Italiana*, in press.
- Tassoul,M.: 1980, *Astroph.J.*, **43**,469.
- Delache, P. and Scherrer, P.H., 1983, *Nature*, 306, 651.

**ACKNOWLEDGMENTS:** The authors would like to acknowledge Dr.R.C. Willson, Jet Propulsion Laboratory, Pasadena, California, USA, who kindly provided us with the excellent data set from the first 280 days of his radiometer ACRIM on the Solar Maximum Mission satellite.





OBSERVATIONS OF LOW-DEGREE MODES FROM THE SOLAR MAXIMUM MISSION  
(Extended Abstract)

Martin Woodard  
Center for Astrophysics and Space Sciences, C-011  
University of California, San Diego  
La Jolla, California 92093

Mean frequencies, amplitudes, and linewidths for the solar 5-min p-mode oscillations of degree 0, 1, and 2 have been obtained from ~280 days of SMM-ACRIM total irradiance data (Woodard and Hudson, 1983). The frequencies are in good agreement with measurements obtained from velocity data, and are given in the Table. The amplitudes of the modes lie along a well defined envelope of power vs. frequency, which peaks at 3.1 mHz and has a width of 0.7 mHz (FWHM). The r.m.s. amplitude of the highest peak in the spectrum ( $n=21$ ,  $l=1$ ) is ~3 ppm of the total flux. The linewidths of the narrowest  $l=0$  modes are ~1  $\mu$ Hz (FWHM). A broad "continuum" of power caused both by solar surface granulation and by instrumental noise interferes with the analysis of 5-min modes. The continuum spectral power in a 1  $\mu$ Hz band near 3 mHz corresponds to an apparent r.m.s. variation of ~0.5 parts per million of the mean solar flux.

These results have been interpreted in terms of a model in which the amplitudes of the separate ( $n\ell m$ ) modes obey a damped harmonic oscillator equation and are excited stochastically by broad-band noise. The full width at half maximum of the implied Lorentzian line profiles,  $\Delta\omega$  (angular frequency), is related to the e-folding decay time,  $\tau$ , of the squared mode amplitude:

$$\tau = 1/\Delta\omega$$

giving  $\tau = \sim 2$  days for the longest-lived  $l=0$  modes. Consistent with this estimate is an upper limit  $\tau = 6$  days inferred from the dispersion in the power of the dominant  $l=0$  modes. To obtain this limit, I assume that the spread in power within the ensemble of peaks in the overall spectrum indicates the time variations in the power of a typical individual mode. The chaotic oscillator model predicts the relation

$$\sigma^2 = 2\tau/T$$

for the fractional variation in power,  $\sigma$ , in an interval of duration  $T$ . Allowance was made for dependence of the mode power upon frequency by

taking the deviation in mode power about a best-fit parabolic envelope.

No fine structure of the type claimed by the Birmingham group ("rotational" splitting) is seen in the  $l=1$  or 2 modes. The  $l=1$  modes appear to be broader than the  $l=0$  modes, an indirect indication of splitting. If, as seems reasonable, the  $m$ -substates of  $l=1$  multiplets have the same width as the  $l=0$  modes, then the observed  $l=1$  widths can be used to measure solar internal rotation. Experiments with simulated data show that the signal-to-noise ratio of the ACRIM data is capable of distinguishing the splitting ( $\sim 0.43 \mu\text{Hz}$  per  $m$ -state) expected from uniform solar rotation at the surface rate from the splitting corresponding to a mean interior rotation rate exceeding twice the surface rate. Precise limits will be the subject of a future paper.

Several additional analyses have been performed. Time variations in the power of individual  $l=0$  modes ( $n=19, \dots, 23$ ) among non-overlapping  $\sim 50$ -day sub-intervals of data have been measured directly and imply a lifetime between 1 and 5 days, in agreement with the 2-day lifetime inferred from linewidth. An attempt was made to measure secular changes in the frequencies of normal modes by dividing the data into  $\sim 100$ -day data sub-strings. The amount of variation is consistent with the random fluctuation expected from the oscillator model of excitation and with the amount of background noise in the data. If the frequencies of the main  $l=0$  and 1 modes vary in unison then the amount of variation which occurred over the  $\sim 10$  month observation span must be significantly less than  $1 \mu\text{Hz}$ . Uncorrelated frequency drifts of this magnitude cannot be discounted, however.

The statistical uncertainty,  $\delta\nu$ , in the frequency estimate from a data string of length  $T$ , computed with the aid of the chaotic oscillator model, is ideally given by

$$(2\pi\delta\nu)^2 = 1/2TT$$

so that for  $\tau = 2$  days and  $T = 300$  days  $\delta\nu = 0.05 \mu\text{Hz}$ . More realistic error estimates which take into account the background noise level of the data imply errors of  $\sim 0.2 \mu\text{Hz}$  for the dominant modes. These error bars are significantly less than the frequency splittings expected from uniform solar rotation and indicate the accuracy to which the solar internal rotation rate can eventually be inferred from  $p$ -mode data. Detailed results concerning frequency and amplitude variations and the method of error analysis will be presented in a future paper.

The author thanks Roger Ulrich for covering his travel expenses, and Hugh Hudson for useful discussions and continuing support.

#### Reference

Woodard, M. and Hudson, H.S., 1983, Nature **305**, p. 589.

Table of frequency centroids ( $\mu\text{Hz}$ )

(Note that the errors quoted here are larger than those in the reference article, as a result of further tests with simulated data.)

	Degree, $l$	0	1	2
Order	$n$			
	17		2559.2 $\pm$ 0.3	2619.7 $\pm$ 0.5
	18	2629.5 $\pm$ 0.5	2693.7 $\pm$ 0.2	2755.0 $\pm$ 0.5
	19	2765.0 $\pm$ 0.4	2828.7 $\pm$ 0.2	2890.0 $\pm$ 0.4
	20	2899.1 $\pm$ 0.3	2963.5 $\pm$ 0.2	3024.5 $\pm$ 0.3
	21	3034.0 $\pm$ 0.2	3098.7 $\pm$ 0.2	3160.0 $\pm$ 0.3
	22	3169.4 $\pm$ 0.2	3233.4 $\pm$ 0.2	3295.2 $\pm$ 0.5
	23	3303.8 $\pm$ 0.3	3369.5 $\pm$ 0.2	3433.2 $\pm$ 0.5
	24	3439.8 $\pm$ 0.3	3505.1 $\pm$ 0.3	
	25		3642.5 $\pm$ 0.5	



## IMPLICATIONS OF OBSERVED FREQUENCIES OF SOLAR P MODES

Jørgen Christensen-Dalsgaard  
HAO/NCAR, Boulder, Colorado, USA and NORDITA,  
Blegdamsvej 17, DK-2100 København Ø, Denmark<sup>1)</sup>

Douglas Gough  
Department of Applied Mathematics and Theoretical  
Physics and Institute of Astronomy, University of  
Cambridge, England

**Abstract:** We present a preliminary comparison of the observed frequencies of 5-min modes reported by Duvall & Harvey in these proceedings with theoretical frequencies for a traditional solar model. The differences between observations and theory can be understood qualitatively in terms of two separate sources of error in the frequency calculation, one near the solar surface and the other at the base of the convection zone. There is no indication of errors in the deep interior of the model.

### 1 Introduction

Duvall & Harvey (these proceedings) present tables of frequencies of individual 5-min p modes of degree  $\ell$  from 1 to 14, as well as frequencies at selected values of  $\ell$  from 20 to 200 obtained by fitting spline functions to ridges in a two-dimensional power spectrum of the data. The accuracy of the former set of frequencies is given as 1  $\mu$ Hz; the latter set is less accurate. These results constitute the largest set of solar oscillation data published to date. Undoubtedly such and even better data will eventually be used for an actual inversion (Gough, these proceedings); however it is of immediate interest to compare them with the frequencies predicted by a traditional solar model. The simple comparison presented here was carried out during the meeting, with the help of a pocket calculator, using a printed table of theoretical frequencies.

Before considering the results it is instructive to discuss briefly the behaviour of the modes of 5-min

#### 1. Present address

oscillation (see also the review by Toomre, these proceedings). A mode with  $\ell > 0$  penetrates to the point where its wave number vector is horizontal, at a radius  $r_t$  given by

$$\frac{r_t}{c(r_t)} = \frac{\sqrt{\ell(\ell+1)}}{2\pi\nu} \quad (1)$$

where  $c$  is the adiabatic sound speed and  $\nu$  is the cyclic frequency of oscillation. Radial modes propagate vertically and penetrate to the centre. The modes considered here have a fairly restricted range of  $\nu$ . Consequently the positions of the turning points  $r_t$  are determined chiefly by  $\ell$  and by the variation of sound speed in the equilibrium model. Plots of  $r_t$  have been given by Duvall & Harvey (1983) and by Christensen-Dalsgaard (1983). For a mode with given  $\ell$  and  $\nu$ , the frequency is sensitive only to the structure of that part of the model which lies outside  $r = r_t$ .

At a given radius  $r$  in the model, the sound propagation time across a horizontal wavelength is

$$t_h = \frac{2\pi r}{c(r) \sqrt{\ell(\ell+1)}} \quad (2)$$

Thus at  $r = r_t$ ,  $t_h$  is equal to the period  $\nu^{-1}$  of oscillation. Further from the centre  $t_h$  exceeds the period, and so there, roughly speaking, there can be no horizontal propagation of information across a wavelength during a period, and the horizontal structure of the mode, as specified by  $\ell$ , has little direct influence on the dynamics.

Thus the eigenfunctions of modes with approximately the same frequency, but different degrees, should be very similar except near the turning points. This has been confirmed by numerical computations.

These simplified arguments suggest how one might recognize the effects on the oscillation frequencies of localized changes in the equilibrium model or in the oscillation calculation. A modification very close to the surface of the model should cause changes in the frequencies that depend on frequency but not on degree. A modification located in a deeper region has no effect on modes with turning points above that region, but it causes changes for modes with turning points in or slightly below it that depend on both frequency and degree. Frequencies of modes

with turning points considerably deeper suffer changes that are determined predominantly by frequency, and only to a lesser extent by degree. These properties are used in the following interpretation of the observations of Duvall & Harvey.

## 2 Results

Fig. 1 illustrates differences  $\Delta\nu$  between frequencies observed by Duvall & Harvey and corresponding adiabatic frequencies computed for Model 1 of Christensen-Dalsgaard (1982). For clarity, frequency differences for modes of the same degree have been connected with continuous (for  $\ell \leq 20$ ) or dashed (for  $\ell > 40$ ) lines.

One is struck by the fact that the modes appear to fall into two distinct groups. Modes of degree  $\ell \leq 20$  all lie within a sloping band with a width of a few  $\mu\text{Hz}$ : here, therefore,  $\Delta\nu$  is almost exclusively a function of  $\nu$ , the variation with  $\ell$  being largely consistent with the scatter in the observed frequencies. Similarly the modes with  $\ell > 40$  have frequency differences within a somewhat broader band with a steeper slope. Here the scatter is larger. There is a slight hint of a weak systematic dependence on  $\ell$ , though that may be the result of observational error, which Duvall & Harvey quote as being  $\pm 10 \mu\text{Hz}$ . In particular, the large excursions between adjacent modes, seen for example for  $\ell = 60$  and  $\ell = 200$ , are almost certainly observational.

## 3 Discussion

The results shown on Fig. 1 may be interpreted in terms of the discussion in Section 1. The fact that  $\Delta\nu$  is predominantly a function of  $\nu$  within each band indicates that for each group of modes the errors in the computation of the frequencies arise in a region that is substantially closer to the surface than are the turning points for all the modes in the group. For the group with  $\ell > 40$  this implies that the sources of error are localized to the outer 5 per cent by radius of the Sun. The difference in the slope between the two bands may be understood in terms of additional sources of error located between the turning points corresponding to  $\ell = 20$  and  $\ell = 40$ , namely at a fractional radius of 0.6 - 0.7. Apparently these partly compensate for the errors at the surface, thereby reducing  $\Delta\nu$  for low degrees. Finally it is worth noticing that the absence of significant variation of  $\Delta\nu$  with  $\ell$  when  $\ell \leq 20$  suggests that to within the accuracy of these observations the deep interior of the Sun is correctly described by a traditional solar model.

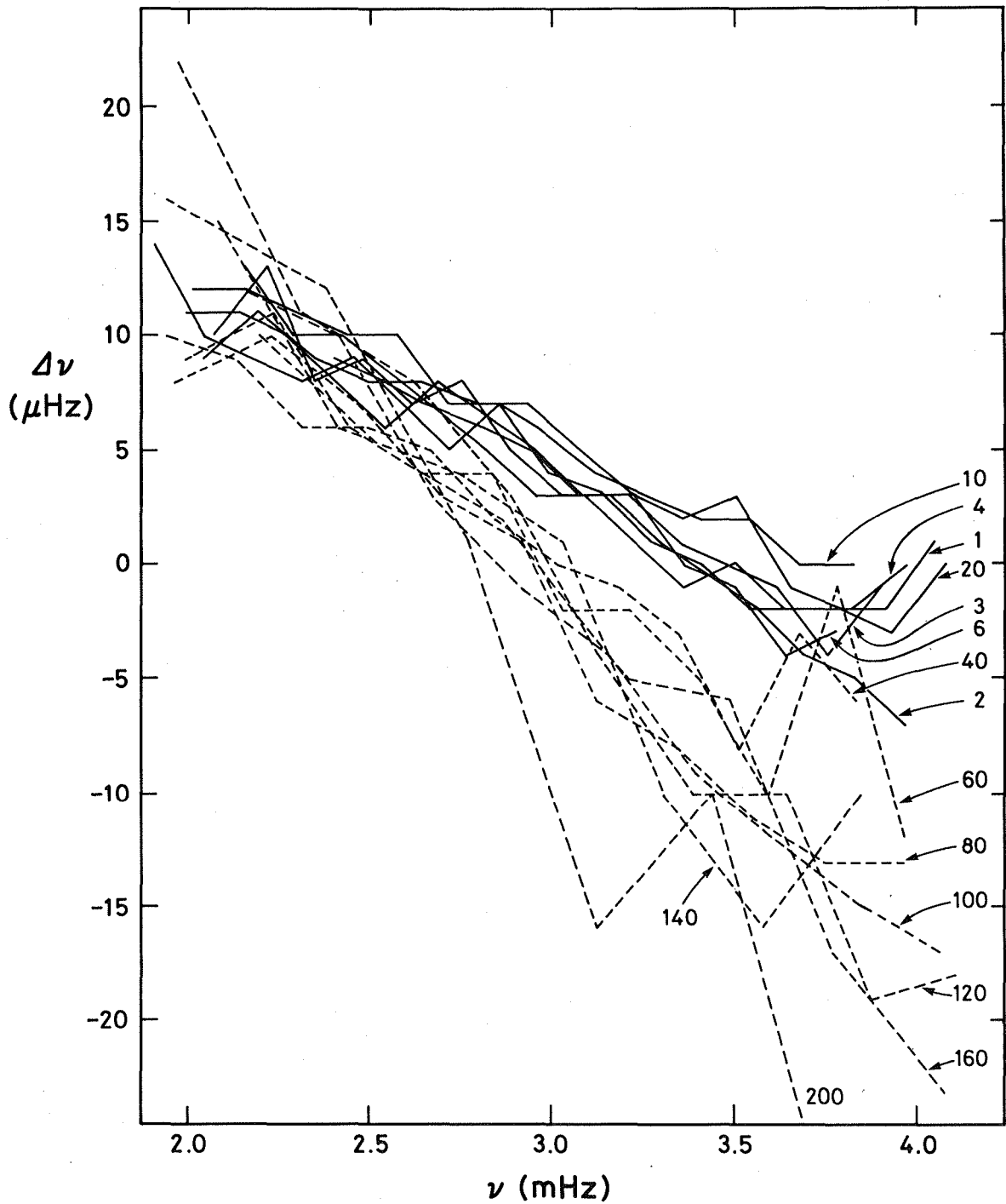


Figure 1. Differences  $\Delta\nu$  between the frequencies observed by Duvall & Harvey (these proceedings) and adiabatic frequencies computed for a traditional solar model:  $\Delta\nu$  is shown as a function of frequency  $\nu$ , and for clarity modes with the same degree have been connected with continuous ( $l \leq 20$ ) or dashed ( $l \geq 40$ ) lines. The values of  $l$  are indicated in the figure.



Tests have shown that numerical errors in the calculation of the equilibrium model cause frequency errors that are smaller than the  $\Delta\nu$  obtained here by a factor of about 4. The numerical errors associated directly with the frequency calculation are quite insignificant (see also Christensen-Dalsgaard 1982). Thus there must be significant errors in the physical assumptions behind the calculation of the frequencies.

Several effects may contribute to the errors near the solar surface. Nonadiabaticity is important only very near the surface, but its effect on the frequencies is probably substantially smaller than the observed  $\Delta\nu$ . Furthermore it has a much stronger dependence on  $\nu$  than that found here. However one cannot exclude the possibility that the excitation of the modes, whatever its source, might cause frequency shifts comparable with the differences observed. Within the framework of adiabatic oscillations, inaccuracies in the equation of state might cause errors in the adiabatic exponent  $\Gamma_1$ , which would affect the oscillation frequencies; it is plausible that such errors would largely be restricted to the outer 5 per cent or so of the model. In addition, the eigenfrequency calculation does not account for turbulent pressure, which may modify the structure of the equilibrium model and affect the dynamics of the oscillations near the surface.

Concerning the errors in the interior of the model, it should be noticed that their inferred location is quite close to the lower boundary of the convection zone. This could indicate an error in the depth of the convection zone, or that the neglect of convective overshoot in the calculation of the solar model has a significant effect on the calculated frequencies.

#### 4 Conclusion

The differences between the frequencies observed by Duvall & Harvey (these proceedings) and those calculated for a traditional solar model (Christensen-Dalsgaard 1982) can be understood in terms of two separate sources of error in the frequency calculation: one located near the surface of the model and the other near the base of the convection zone. The former might be caused by, for example, the excitation of the oscillations, errors in the equation of state or effects of turbulent pressure; the latter could result from an error in the depth of the convection zone, or from the neglect of convective overshoot. The results give no indication of errors in the energy-generating part of the model; it remains to be seen whether that is consistent with the partial mixing that Berthomieu, Provost & Schatzman

(1984) have proposed to account for the period spacing observed for high-order g modes.

The present analysis is clearly preliminary, and largely qualitative. To verify the existence of two separate sources of error more and better observations are needed, in particular for modes with degree higher than 20. In addition it would be useful to study the effects on the frequencies of specific changes in the equilibrium model, attempting to reproduce the features of the observed  $\Delta\nu$  and to place limits on the set of models that are consistent with the observations. The effects on the frequencies of mode excitation and turbulent pressure must also be studied.

A formal inversion of the data will be attempted, to determine the change in the model required to bring it into agreement with the observed frequencies. This, however, is not completely straightforward owing to uncertainties in the physics of the oscillations close to the solar surface. It is unlikely that adequate theories of turbulent pressure and mode excitation will be formulated in the near future. Therefore the inversion procedure discussed by Gough (these proceedings) will probably need to be extended to include a component of the frequency difference that accounts for these uncertainties.

NCAR is sponsored by the National Science Foundation

### References

- Berthomieu, G., Provost, J. & Schatzman, E., 1984. Mem. della Societa Astronomica Italiana, in press.  
Christensen-Dalsgaard, J., 1982. Mon. Not. R. astr. Soc., **199**, 735.  
Christensen-Dalsgaard, J., 1983. Advances Space Res., **2**, 11.  
Duvall, T. L. & Harvey, J. W., 1983. Nature, **302**, 24.

PART 3

ANALYSIS AND INTERPRETATION OF OBSERVATIONS, SPATIAL FILTERS,  
SEEING AND DATA GAP FILLING



OBSERVED SPATIAL PROPERTIES OF THE SOLAR EIGENFUNCTIONS AND  
THE IMPLICATIONS FOR THE EXISTENCE OF RESOLVED MULTIPLETS

Thomas P. Caudell  
SCLERA, Department of Physics  
University of Arizona, Tucson, AZ, USA

Henry A. Hill  
Professor, Department of Physics  
University of Arizona, Tucson, AZ, USA

and

Randall J. Bos  
SCLERA, Department of Physics  
University of Arizona, Tucson, AZ, USA

INTRODUCTION

Solar oscillations are manifested in the solar atmosphere as spatial and temporal perturbations in the local thermodynamical and mechanical properties. When measuring the solar radius/diameter, these perturbations enter the observation through changes in the radiative source function and opacity at the extreme limb. When compared to the disk center, the observable portion of these perturbations is changed in spatial character by projection effects and oblique optical-depth geometry. The time-varying solar radius signal at SCLERA<sup>1</sup> is produced by an edge definition sensitive to the resultant changes in the spatial shape of the limb intensity profile. The object of this study is to interlace theory and observation in an attempt to further determine the shape and properties of the limb signals which display these global solar oscillations. The following sections review the observations and discuss the methods and results of this study.

---

<sup>1</sup>SCLERA is an acronym for the Santa Catalina Laboratory for Experimental Relativity by Astrometry jointly operated by the University of Arizona and Wesleyan University.

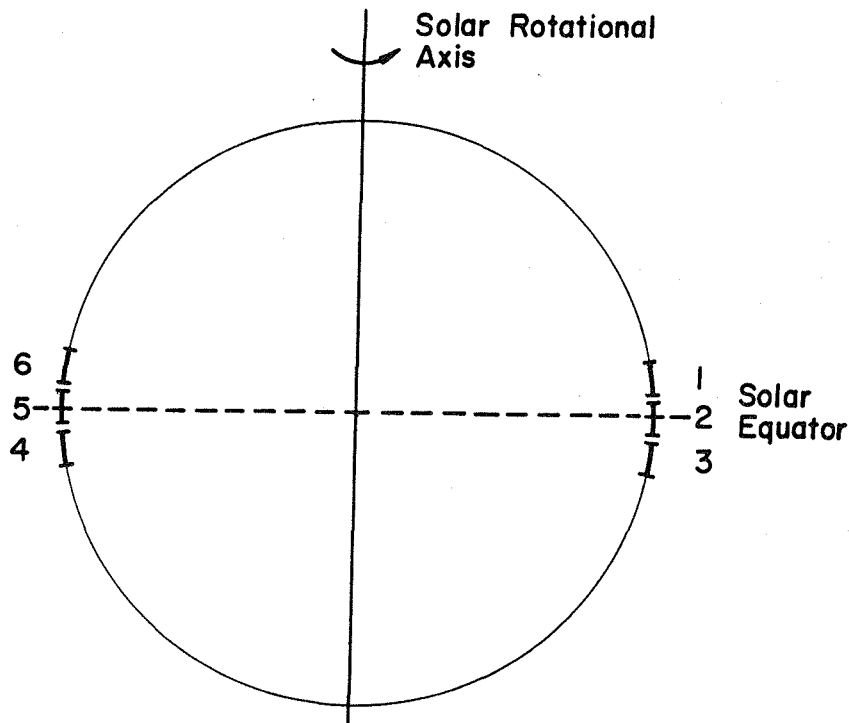


Fig. 1. The detector geometry for the 1979 SCLERA data set. The numbers label the locations of the six limb intensity photodetectors.

#### CHARACTERISTICS OF THE 1979 SCLERA DATA

The data considered in this analysis were collected in the summer of 1979 using the SCLERA instrument. Both the instrument and the details of the data analysis have been described in other works (Stebbins 1975, Bos 1982); therefore, only a brief summary of the salient points will be given here.

Over a period of 41 days, time strings were recorded of intensity profiles at the extreme limb; the measurements were taken at six position angles near the solar equator. The detector geometry is shown in Figure 1. Off-line, the finite Fourier transform definition (FFTD) was applied twice to each profile with window sizes of  $27''2$  and  $6''8$  [see Hill, Stebbins and Oleson (1975) for a detailed description of the FFTD]. Differences were then taken between the two defined radii to form six time strings of radius functionals primarily sensitive to changes in limb profile shape. Finally, linked baseline temporal Fourier transforms were performed on the six time strings over the entire observing period.

By linearly combining the six transforms, different symmetry conditions can be enforced, controlling the parity of the spatial part of the oscillation about the solar equator or the north-south line. Mathematically, the solutions which describe an oscillation are products of a radial eigenfunction and a spherical harmonic, with order numbers  $n$  and  $(\ell, m)$  respectively. Setting the parity amounts to assigning the  $\ell$  and  $m$  order numbers to be even or odd. This property somewhat reduces the spectral peak density in a given linear combination and aids in the identification of modes.

Two linear combinations are used in this analysis;  $F_1$  and  $F_2$  refer to the sum and difference, respectively, of the two diametrically opposite groups of limb signals. Each group consists of the sum of three individual radius functionals. Since the Fourier transform is a linear operator, combining the individual transforms is equivalent to combining the radius functional time strings; these procedures therefore may be used interchangeably. The properties of the observed signal combinations may be quantified by first assuming the spatial and temporal form of the intensity perturbation at the limb to be

$$I'_{\ell m}(u, t) = I'_{\ell 0}(u) \cdot \exp \left\{ i (\phi_m + m\phi + \omega t) \right\} \quad (1)$$

where  $u$  is the distance from the nominal disk edge,  $I'_{\ell 0}(u)$  is the  $m = 0$ ,  $t = 0$  radial perturbation,  $\phi$  is the standard angle in a spherical coordinate system,  $\phi_m$  is the initial phase angle of the particular state in the  $\phi$  direction, and  $\omega$  is the temporal angular frequency. We define the two combinations to be

$$\begin{aligned} F_1 &= \text{FFTD} \left[ f_1(\ell, m, \phi_m, \phi, \omega, t) \right] \\ F_2 &= \text{FFTD} \left[ f_2(\ell, m, \phi_m, \phi, \omega, t) \right] \end{aligned} \quad (2)$$

where

$$\begin{aligned} f_1 &= \left[ \frac{1 + (-1)^{\ell+m}}{2} \right] \left[ \text{Re } I'_{\ell 0} \cos(\phi_m + \omega t) - \text{Im } I'_{\ell 0} \sin(\phi_m + \omega t) \right] \\ &\quad \left\{ (-1)^{m/2} \left[ \frac{1 + (-1)^m}{2} \right] \cos m \left( \frac{\pi}{2} - \phi_+ \right) \right. \\ &\quad \left. + (-1)^{(m-1)/2} \left[ \frac{1 - (-1)^m}{2} \right] \sin m \left( \frac{\pi}{2} - \phi_+ \right) \right\} \end{aligned} \quad (3)$$

$$\begin{aligned}
f_2 = & \left[ \frac{1 + (-1)^{\ell+m}}{2} \right] \left[ \text{Re } I'_{\ell 0} \sin(\phi_m + \omega t) + \text{Im } I'_{\ell 0} \cos(\phi_m + \omega t) \right] \\
& \left\{ (-1)^{m/2} \left[ \frac{1 + (-1)^m}{2} \right] \sin m \left( \frac{\pi}{2} - \phi_+ \right) \right. \\
& \left. - (-1)^{(m-1)/2} \left[ \frac{1 - (-1)^m}{2} \right] \cos m \left( \frac{\pi}{2} - \phi_+ \right) \right\}, \tag{3}
\end{aligned}$$

and  $\phi_+$  is  $\phi$  measured positively from the disk edge toward the center. The notation FFTD [ ] symbolizes the conversion of changes in the limb darkening function into radius functionals defined as differences between two FFTD edges;  $F_1$  and  $F_2$  are in units of distance.

Specific information about  $I'_{\ell m}$  can be obtained by taking a ratio between these two quantities at the assigned temporal frequency for a given mode of oscillation. This procedure allows the testing of the spherical harmonic representation of the eigenfunction, particularly the  $e^{im\phi}$  portion. An additional consistency test of the symmetry properties may thus be made. Measurement of the radial shape and phase of the intensity perturbation is also possible. Further, by taking the ratio of  $F_1$  and  $F_2$ , cancellation is effected of differences in modal amplitude due either to excitation or spatial filtering of the  $P_{\ell m}(\theta, \phi)$ ; comparison between modes is then possible. In practice, extraction of the above information requires that the ratio be studied as a function of azimuthal order number  $m$ . This method implicitly uses the frequency splitting of the states into resolved multiplets.

In the work of Hill and Bos (1984), a series of 32 rotationally split multiplets has been identified; these states vary in degree  $\ell$  from 2 to 22 and in nodal index  $n$  from 1 to 3. The frequencies range between 450 and 800  $\mu\text{Hz}$ , implying the  $p$ -mode assignment. At the time of this writing, four of these multiplets were available for ratio analysis (see Table 1). Figures 2 and 3 display separate averages of the real and imaginary parts of the ratio, with negative  $m$  states having been reflected through the origin and averaged with positive  $m$ . Given the size of the error bars, significant structure is present in both the real and imaginary portions. The object of this analysis is to invert the structure evident in this data and extract the properties of the source intensity perturbation.



TABLE 1. Four Multiplets Used in Inversion Analysis

$P_{n,1} : P_{1,10} P_{1,14} P_{1,22} P_{3,4}$

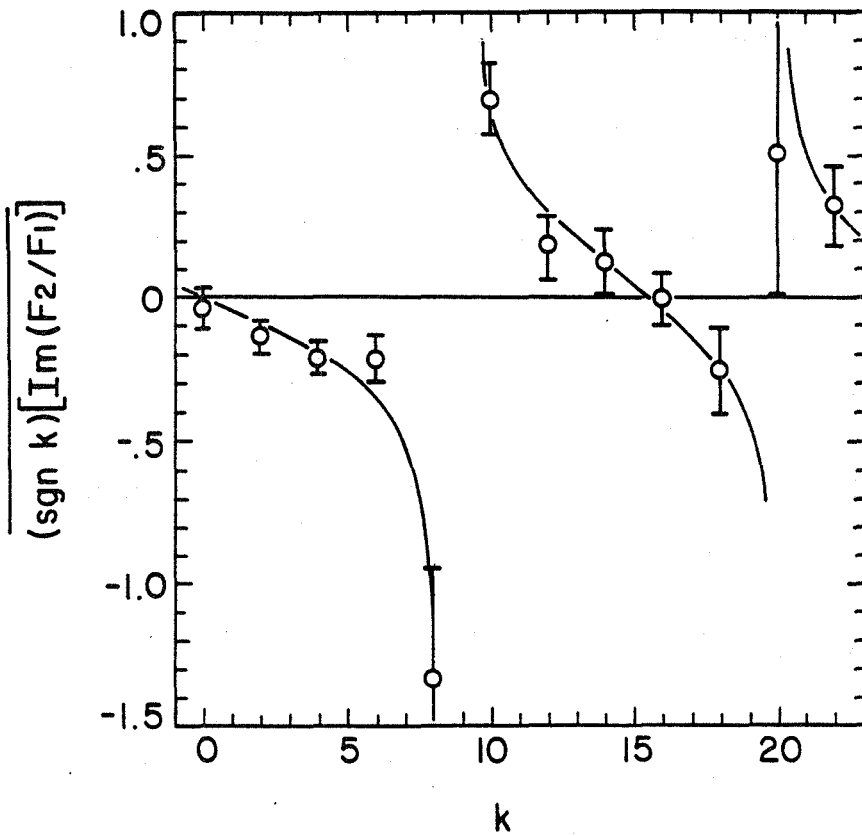


Figure 2. Imaginary part of ratio  $(F_2/F_1)$  averaged in  $m$  for four multiplets. The smooth curve is not a fit.

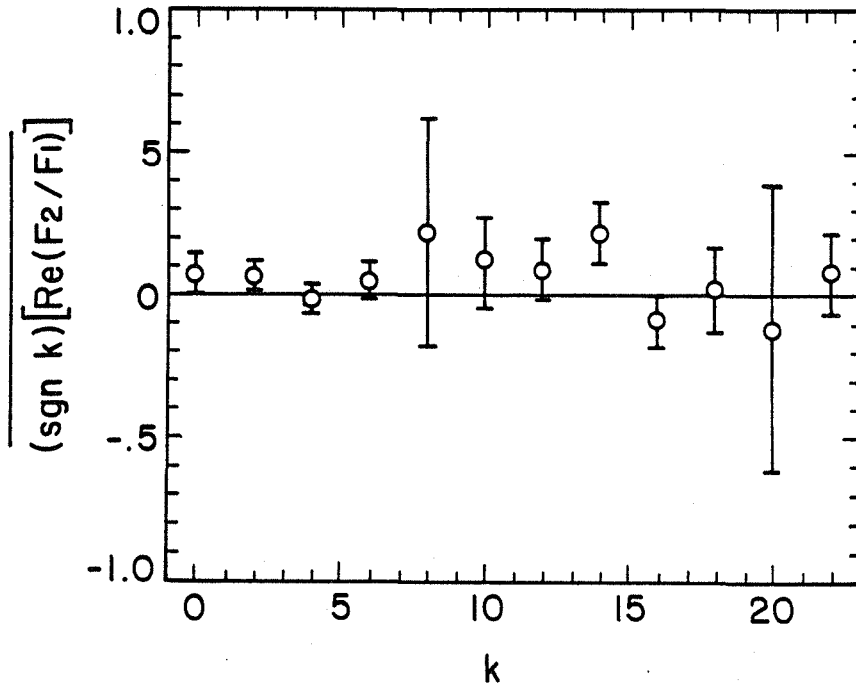


Figure 3. Real part of ratio  $(F_2/F_1)$  averaged in  $m$  for four multiplets.

#### RESULTS OF THE INVERSION TECHNIQUE

A simple representation is needed for the initial inversion. The present work assumes that the intensity perturbation is all of one spatial phase; this phase is also chosen so that  $I'_{\ell_0}(u)$  is real. The radial form of  $I'_{\ell_0}(u)$  is expanded in a generalized power series where a minimum number of individual components are selected to give the best fit to the ratio data. Hence,

$$I'_{\ell_0 m}(u) = \left( \sum_{\alpha} a_{\alpha} g_{\alpha}(u) \right) \left( \cos m\phi(u) + i \sin m\phi(u) \right). \quad (4)$$

The  $a_{\alpha}$  will be determined by the fitting procedure, the  $g_{\alpha}(u)$  are members of a subset of all monomials of the form

$$u^{\pm j/k} \quad (\text{i.e., } u, u^{\pm}, u^{-1}, u^{-1/2}, \text{ etc.}), \quad (5)$$

$\alpha$  is the pointer into this subset, and

$$\phi(u) = \cos^{-1} \left( 1 + \frac{u}{R_{\theta}} \right)$$

where  $R_{\theta}$  is the radius of the sun. Again,  $u$  is a linear distance measure on the solar disk, with the origin at the nominal edge and defined positively outwards. Note that, although the coefficients  $a_{\alpha}$  will generally be complex, they are real in this simplification.

Applying the FFTD operation to the individual components of this representation, we define two new functions

$$\psi_{\alpha}(m) = \text{FFTD} \left[ g_{\alpha}(u) \cdot \sin m\phi(u) \right] \quad (6)$$

and

$$\beta_{\alpha}(m) = \text{FFTD} \left[ g_{\alpha}(u) \cdot \cos m\phi(u) \right]$$

where  $\psi_{\alpha}(m)$  and  $\beta_{\alpha}(m)$  are units of distance and  $u$  has been removed by integration in the  $\alpha$  FFTD procedure. Equation (2) shows the relationship between  $F_1$  and  $\psi_{\alpha}(m)$  and between  $F_2$  and  $\beta_{\alpha}(m)$ . The form of the ratio discussed earlier is now defined to be

$$R(m) = F_2/F_1 = \frac{\sum_{\alpha} a_{\alpha} \psi_{\alpha}(m)}{\sum_{\alpha} a_{\alpha} \beta_{\alpha}(m)}, \quad (7)$$

where the temporal dependence and amplitude cancel. A fitting procedure must now be applied to determine the  $a_{\alpha}$ . Clearly, without loss of generality we may set  $a_0 = 1$  independent of the form of  $g_0(u)$ . Several fitting procedures have been developed to handle the general problem of complex coefficients. A simple approach was chosen here to illustrate the sensitivity of this method. In future work, a general technique will be used to invert the more complete set of ratio data including all 32 multiplets.

In the simple approach, the subset of  $g_{\alpha}(u)$  used is  $u^0, u^{+1}, u^{+1/2}$  necessitating the determination of only two free parameters,  $a_1$  and  $a_2$ . The coefficient  $a_0$  is set to unity as discussed above. Requiring  $R(m)$  to be infinite at  $m = 9$  and zero at  $m = 16$ , values of  $a_1 = -0.0709$  and  $a_2 = -0.5664$  are obtained. A continuous evaluation of  $R(m)$  is given in Figure 4, and Figure 5 shows a reconstruction of  $I'_{\ell_0}(u)$  for this model. For comparison,  $R(m)$  is calculated and plotted for a constant  $I'_{\ell_0}(u)$  (i.e.,  $a_0 = 1, a_1 = 0 = a_2$ ); these results appear in Figure 6.

Comparing the results for the simple model to the ratios averaged over four multiplets in Figure 3, we see that the general structures are quite similar, notwithstanding the difference in overall magnitude. Conversely, the data and Model 2 show little similarity.

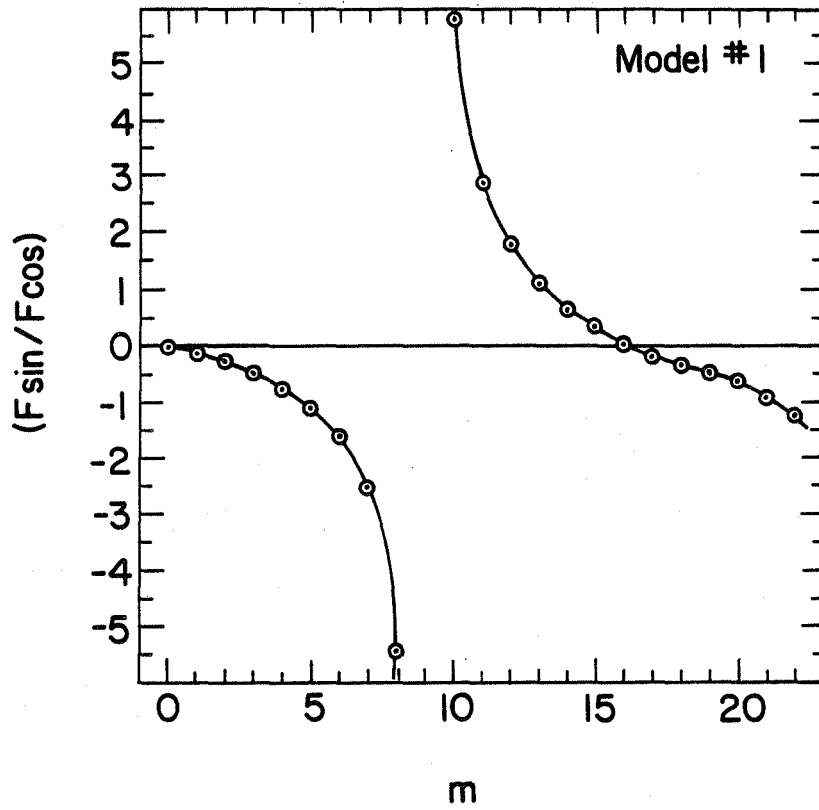


Figure 4. Resultant ratio curve when a singularity is imposed at  $m = 9$  and a zero at  $m = 6$ .

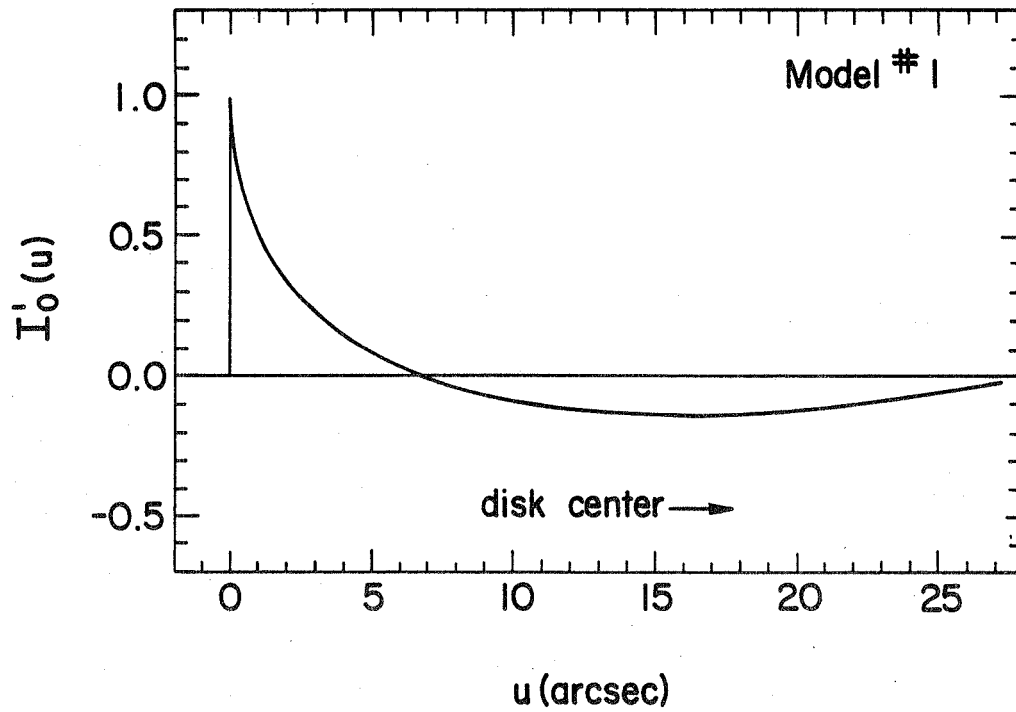


Figure 5. Reconstruction of intensity perturbation fit which gives ratios in qualitative agreement with data.

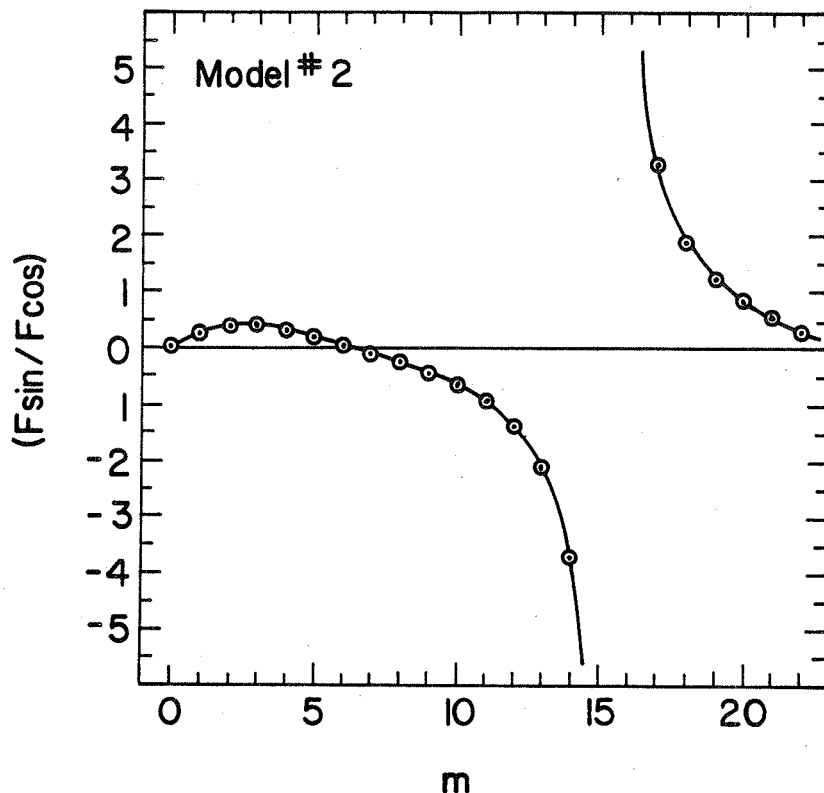


Figure 6. Resultant ratio curve for a spatially constant  $I'_{\lambda 0}(m)$ .

#### CONCLUSION

The simple example presented above serves two purposes. First, it illustrates the potential sensitivity with which radiative processes in the solar atmosphere may be probed. A comparison of Figures 4 and 6 graphically demonstrates this aspect. Second, the shape of the intensity perturbation at the extreme limb must be peaked in the outermost layers of the solar atmosphere, a result consistent with previous work at SCLERA (Hill and Caudell 1979; Knapp et al. 1980). The model used in this work is a simplified one and was applied only to a subset of the data; therefore, the present results must be considered preliminary.

This work was partially supported by the Air Force Office of Scientific Research and the National Science Foundation Division of Atmospheric Sciences.

## REFERENCES

1. Bos, R. J., Ph.D. Thesis, University of Arizona (1982).
2. Hill, H. A. and Bos, R. J., In preparation (1984).
3. Hill, H. A. and Caudell, T. P., Mon. Not. R. Astro. Soc. 186, 327 (1979).
4. Hill, H. A., Stebbins, R. T. and Oleson, J. R., Ap. J. 200, 484 (1975).
5. Knapp, J., Hill, H. A. and Caudell, T. P., Nonradial and Nonlinear Stellar Pulsation, Lecture Notes in Physics No. 125 (Springer - Verlag, Berlin, 1980).
6. Stebbins, R. T., Ph.D. Thesis, University of Colorado (1975).





## OPTIMIZED RESPONSE FUNCTIONS FOR TWO-DIMENSIONAL OBSERVATIONS OF SOLAR OSCILLATIONS

Jørgen Christensen-Dalsgaard  
HAO/NCAR<sup>1)</sup>, Boulder, Colorado, USA and NORDITA,  
Blegdamsvej 17, DK-2100 København Ø, Denmark<sup>2)</sup>

**Abstract:** Reasonably realistic response functions are calculated for solar oscillation observations made in Doppler velocity which is averaged over a grid of 30" x 30" pixels. In a simulation of an analysis scheme proposed for the HAO/SPO Fourier Tachometer the responses for the individual pixels are combined using Chebychev weighting functions. It is shown how a technique developed for geophysical inversion may be used to find linear combinations, concentrated within a fairly narrow range of  $\ell$ - and  $m$ -values, of these transformed responses. The extension of this analysis to bidirectional observations is discussed.

### 1 Introduction

Before observations of solar oscillations can be used for helioseismology, the individual modes of oscillation that are present in the timestring must be isolated and identified. This, however, is greatly complicated by the extreme richness of the spectrum of solar oscillations.

As discussed in the review by Toomre (these proceedings) the structure over spherical surfaces of a mode of oscillation is described by a spherical harmonic  $Y_{\ell}^m(\theta, \phi)$ , which is characterized by its degree  $\ell$ , measuring the total wave number (or, equivalently, the total number of nodal lines on the surface) and its azimuthal order  $m$ , that gives the number of nodal meridians. For each value of  $\ell$  and  $m$  the Sun can support a large number of modes, which differ in their frequencies and in the radial variation of their eigenfunctions, and which are characterized by their radial orders  $n$ . When rotation is neglected the frequencies are

1. NCAR is sponsored by the National Science Foundation
2. Present address

independent of  $m$ : in the presence of rotation this degeneracy is lifted, and so to each pair  $(n, \ell)$  there correspond  $2\ell+1$  different frequencies, distinguished by the value of  $m$ .

In the absence of an adequate theory for the excitation of solar oscillations there is no way to predict a priori which of these possible modes will be present. However it appears from the observations that the excitation mechanisms are generally broad-band, so that a substantial proportion of the modes in a given region of the spectrum is excited. In particular, in the 5 min region the average amplitude per mode depends only weakly on  $\ell$  (Christensen-Dalsgaard & Gough 1982); here most modes in the frequency range between about 2 and 5 mHz are excited to sufficiently high amplitudes to permit their detection in present observations.

An observation made at a single point on the disk of the Sun is equally sensitive to all modes: although power spectrum analysis in time provides separation between modes in frequency, the spectra are far too complex to permit isolation of the individual modes. If instead the signal is averaged over an aperture of characteristic size  $d$  (in units of the solar radius), the sensitivity falls off for  $\ell > d^{-1}$  roughly like  $(\ell d)^{-1}$ , and so here the number of  $\ell$ -values contributing to the observations is  $O(d^{-1})$ . However unless  $d$  is of order unity, so that the aperture covers a large fraction of the disk of the Sun, the data still contain too many modes to permit their isolation.

Thus more sophisticated schemes are needed. Observations of 5 min modes of high degree have generally averaged the signal in the North-South direction: in this way modes with  $m \approx \ell$  are selected. From the resulting dataset in one spatial dimension partial modal isolation in  $m$  can then be accomplished by Fourier transform in the azimuthal angle  $\phi$ , possibly neglecting the curvature of the Sun (e.g. Rhodes, Ulrich & Simon 1977; Deubner, Ulrich & Rhodes 1979; Hill, Toomre & November 1983). Similarly Duvall & Harvey (1983: these proceedings) made averages in the East-West direction to isolate modes with  $m \approx 0$  and performed a Legendre transform in the North-South direction: in this way they were able to isolate quite effectively modes with  $\ell$  between 1 and about 140.

Limb measurements (e.g. Brown, Stebbins & Hill 1978; Yerle 1981; Bos & Hill 1983) constitute a distinct type of 1-dimensional observations. Bos & Hill obtained some discrimination between modes by combining observations made at 6 different positions on the solar limb: however the density of peaks in the resulting spectrum was still, in the most favourable part of the spectrum, close to the limit of

the frequency resolution corresponding to their 41 day time string (cf Christensen-Dalsgaard 1982). Gough & Latour (1984) recently showed how the Backus & Gilbert (1970) technique of optimal averages could be modified to set up combinations of large numbers of observations at the limb so as to obtain optimum isolation of the modes. They found that little separation between modes with the same value of  $\ell - m$  could be obtained; this is a general difficulty associated with limb observation.

To get a more complete modal separation in both  $\ell$  and  $m$ , observations in both spatial dimensions must be used. If observations covering the entire solar surface were possible, the modes corresponding to a given  $(\ell, m)$  could be completely isolated by using the orthogonality of the spherical harmonics. This is not possible with observations from only one direction, or, as in the original SDO proposal for satellite observations of solar oscillations, from two directions in the plane of the ecliptic. The present paper considers how to set up such observations to obtain the optimal isolation.

In considering the criteria for adequate modal isolation, one must take into account the possibility of more or less accidental degeneracy between the frequencies of modes of different  $\ell$ . Ulrich has pointed out that for degrees from about 4 to about 12 there is a fairly large number of cases where  $\nu_{n,\ell,m} \approx \nu_{n-1,\ell+3,m'}$  for  $|m - m'| < 3$  (see also Mihalas, Christensen-Dalsgaard & Brown, these proceedings); here  $\nu_{n,\ell,m}$  is the cyclic frequency of the mode  $(n, \ell, m)$ . If a given observing scheme with maximum sensitivity at  $\ell = \ell_c$  has significant sensitivity at  $\ell = \ell_c \pm 3$ , the modes at  $\ell = \ell_c$  may therefore be affected significantly by interference; this can result in difficulties in mode identification, or in apparent changes in the inferred frequencies or amplitudes (Loumos & Deeming 1978; Christensen-Dalsgaard & Gough 1982). Thus a natural goal for the observing scheme is to limit the sensitivity to modes with  $|\ell - \ell_c| \leq 2$ , for  $\ell_c = 4 - 12$ ; in this way one ensures that the modes at the maximum in sensitivity are unaffected by interference. At higher degrees similar degeneracies occur between modes with a separation in  $\ell$  of 4 or more; but this is less likely to cause problems.

In principle the observing scheme could be set up by determining directly the weights to be applied to each pixel in such a way that the modal isolation is optimized; this would correspond to the approach used by Gough & Latour (1984). This is possible in the two-dimensional case provided the resolution is sufficiently coarse. However with more than a few hundred pixels the computational effort becomes excessive. In this case one must make an initial

analysis, with a preliminary set of weights, possibly based to some extent on observational convenience. The results of this preliminary analysis can then be combined by using the Backus & Gilbert technique to optimize the modal isolation. As an example of this we have considered combinations of observations for a scheme using Chebychev weighting, proposed by T. M. Brown (private communication). Here the optimization resulted in a considerable improvement of the separation of the modes.

We have also applied the Backus & Gilbert technique to form optimal combinations of observations made from two separate directions in the plane of the solar equator. This provides a longer base line in  $\phi$ , and correspondingly a better resolution in  $m$ ; there is no significant direct improvement in the resolution in  $\theta$ , in particular for small  $m$ .

A basic assumption is that we can characterize the sensitivity of the observations to a specific mode by a spatial response function<sup>3)</sup> (e.g. Dziembowski 1977; Hill 1978; Christensen-Dalsgaard & Gough 1982), which may be calculated by considering the mode in isolation, neglecting other velocity fields in the solar atmosphere. The combined observed signal then consists of the sum over all the modes present of their apparent velocities, weighted by the response, with the addition of the effects of the other velocity fields which contribute to the noise in the data. This would be true for a strictly linear observing scheme, such as the Fourier Tachometer (Brown, these proceedings); however for a scheme that is not linear, such as the magneto-optical filter (Cacciani & Rhodes, these proceedings) the other velocity fields directly affect the sensitivity of the observations. Brookes, Isaak & van der Raay (1978) discussed the effect of solar rotation on the sensitivity of whole disk resonance cell observations; the effects of nonlinearity on spatially resolved observations have yet to be analyzed.

---

3. Starting with Hill (1978) the term 'spatial filter function' has been employed to denote the sensitivity of observations of solar oscillations to specific modes. This, however, is in conflict with the well-established practice of using 'filter' for the weights applied to a signal. Thus we propose henceforth to use 'spatial response function' instead of 'spatial filter function'.

## 2 Mathematical preliminaries

For a single linear, undamped mode of oscillation the velocity field on the solar surface can be written, in spherical polar coordinates  $(\theta, \phi)$

$$\begin{aligned} \vec{v}_{n,\ell,m}(\theta, \phi; t) = & (-1)^m c_{\ell,m} \operatorname{Re} \left\{ v_{n,\ell,m} P_{\ell}^m(\cos\theta) \vec{a}_r \right. \\ & + \frac{U_{n,\ell,m}}{\sqrt{\ell(\ell+1)}} \left[ \frac{dP_{\ell}^m(\cos\theta)}{d\theta} \vec{a}_{\theta} \right. \\ & \left. \left. + \frac{im}{\sin\theta} P_{\ell}^m(\cos\theta) \vec{a}_{\phi} \right] e^{i(m\phi - \omega_{n,\ell,m} t + \epsilon_{n,\ell,m})} \right\}, \end{aligned} \quad (2.1)$$

where  $P_{\ell}^m$  is the associated Legendre function.<sup>4)</sup> The factor  $(-1)^m$  was introduced for consistency with the definition of the spherical harmonics. Furthermore  $c_{\ell,m} > 0$  is a normalization constant chosen such that

$$c_{\ell,m}^2 \int_{-1}^1 P_{\ell}^m(x)^2 dx = \frac{4}{1 + \delta_{m0}}, \quad (2.2)$$

where  $\delta$  is the Kronecker delta; then  $|V_{n,\ell,m}|$  and  $|U_{n,\ell,m}|$  are the rms over the solar surface of the vertical and horizontal components of the velocity. We neglect the variation in  $V_{n,\ell,m}$  and  $U_{n,\ell,m}$  with height in the atmosphere, and assume that  $\omega_{n,\ell,m} > 0$ . Finally  $\omega_{n,\ell,m} =$

---

4. No unique sign convention has apparently been established for the associated Legendre functions. Here we follow Abramowitz & Stegun (1965), so that for  $m > 0$

$$P_{\ell}^m(x) = (-1)^m (1-x^2)^{m/2} \frac{d^m P_{\ell}(x)}{dx^m}$$

and

$$P_{\ell}^{-m}(x) = (-1)^m \frac{(\ell-m)!}{(\ell+m)!} P_{\ell}^m(x)$$

where  $P_{\ell}(x)$  is a Legendre polynomial.

$2\pi\nu_{n,\ell,m}$  is the angular oscillation frequency, which is real.

The ratio  $U_{n,\ell,m}/V_{n,\ell,m}$  is determined by the linear eigenfunction. Christensen-Dalsgaard & Gough (1982), equations (4.5) - (4.7) give a simple approximation to this ratio, according to which

$$\frac{U_{n,\ell,m}}{V_{n,\ell,m}} \lesssim \sqrt{\ell(\ell+1)} \left( \frac{v_a}{v_{n,\ell,m}} \right)^2 \approx \left( \frac{v_{0,\ell,0}}{v_{n,\ell,m}} \right)^2, \quad (2.3)$$

with approximate equality when  $\nu_{n,\ell,m}$  is substantially smaller than the acoustical cut-off frequency in the solar atmosphere. Here  $\nu_a^2 = GM/(4\pi^2 R^3)$  (where  $G$  is the gravitational constant and  $M$  and  $R$  the mass and radius of the Sun), corresponding to  $\nu_a \approx 0.1$  mHz. In the last approximate equality in equation (2.3) we used that for moderate or large  $\ell$ ,  $\ell^{1/2}\nu_a \approx \nu_{0,\ell,0}$ , the frequency of the f mode (e.g. Gough 1980). We are here principally concerned with modes with  $\nu$  greater than 2 mHz and  $\ell$  less than 50. Here  $U_{n,\ell,m}/V_{n,\ell,m} < 0.1$ . This justifies that we neglect the horizontal component of the velocity in the following. It could, however, be included without any fundamental difficulty.

We assume that the observed spectral line shift is simply caused by the Doppler shift corresponding to the line-of-sight component of the velocity vector in equation (2.1). For simplicity we choose a coordinate system with polar axis in the plane of the sky, and take the line of sight to correspond to the  $\phi = 0$  axis. With this choice the line-of-sight velocity for a single mode is

$$\begin{aligned} & v_{D;n,\ell,m}(\theta, \varphi; t) \\ &= V_{n,\ell,m} (-1)^m c_{\ell,m} P_{\ell}^m(\cos\theta) \\ & \times \cos(m\varphi - \omega_{n,\ell,m} t + \epsilon_{n,\ell,m}) \sin\theta \cos\varphi, \end{aligned} \quad (2.4)$$

and the combined local line shift is given by

$$v_D(\theta, \phi; t) = \sum_{n, \ell, m} v_{D; n, \ell, m}(\theta, \phi; t) . \quad (2.5)$$

The frequencies  $\nu_{n, \ell, m}$  are in fact only well-defined when the polar axis coincides with the axis of symmetry (provided it exists: cf Isaak 1982, Gough 1982, Dicke 1982), and we assume that this is the axis of surface rotation. Thus equation (2.4) is only valid twice a year, when the rotation axis is in the plane of the sky. Otherwise the inclination of the rotation axis, if not corrected for, causes a mixing of modes with the same degree. As discussed in Appendix A this may be calculated using the transformation formulae for spherical harmonics (e.g. Edmonds 1960): for  $\delta = 10$  the mixing involves about 5 values of  $m$  at the maximum inclination of  $7.25^\circ$ .

The quantity resulting from the observation can in general be represented as

$$v_{\text{obs}}(t; \mu) = \int_A W(\theta, \phi; \mu) v_D(\theta, \phi; t) dA \quad (2.6)$$

where  $W(\theta, \phi; \mu)$  is a set of weight functions, labelled by  $\mu$ , and the integration is over the area  $A$  of the solar disk: we assume that the weight functions are real (the possibility of complex weights is briefly discussed in Section 4, below). From equations (2.4) and (2.5) we obtain

$$v_{\text{obs}}(t; \mu) = \sum_{n, \ell, m} V_{n, \ell, m} \psi_m \left[ \tilde{S}_{\ell, |m|}(\mu) \cos(\omega_{n, \ell, m} t - \epsilon_{n, \ell, m}) + \zeta_m \tilde{S}_{\ell, -|m|}(\mu) \sin(\omega_{n, \ell, m} t - \epsilon_{n, \ell, m}) \right] , \quad (2.7)$$

where the spatial response function  $\tilde{S}_{\ell, m}$  is defined by

$$\tilde{S}_{\ell, m}(\mu) = (-1)^m \int_A W(\theta, \varphi; \mu) c_{\ell, m} P_{\ell}^m(\cos\theta) \cos m\varphi \sin\theta \cos\varphi dA, \quad \text{for } m \geq 0 \quad (2.8)$$

$$\tilde{S}_{\ell, -m}(\mu) = (-1)^m \int_A W(\theta, \varphi; \mu) c_{\ell, m} P_{\ell}^m(\cos\theta) \sin m\varphi \sin\theta \cos\varphi dA, \quad \text{for } m > 0.$$

Furthermore the sign factors  $\zeta_m$  and  $\psi_m$  are defined by

$$\zeta_m = \begin{cases} 1 & \text{for } m > 0 \\ 0 & \text{for } m = 0 \\ -1 & \text{for } m < 0 \end{cases} \quad (2.9)$$

and  $\psi_m = 1$  for  $m \geq 0$ ,  $\psi_m = (-1)^m$  for  $m < 0$ . The added complication of having to introduce  $\psi_m$  stems from the sign convention for  $P_{\ell}^m$  (cf. footnote 4).

Finally, assuming an infinitely long observing sequence, the temporal Fourier transform of the observed signal is

$$\begin{aligned} \tilde{v}_{\text{obs}}(\omega; \mu) &= \frac{1}{2\pi} \int v_{\text{obs}}(t; \mu) e^{i\omega t} dt \\ &= \frac{1}{2} \sum_{n, \ell, m} v_{n, \ell, m} \psi_m \left\{ \left[ \tilde{S}_{\ell, |m|}(\mu) + i\zeta_m \tilde{S}_{\ell, -|m|}(\mu) \right] \times \right. \\ &\quad \left. e^{i\epsilon_{n, \ell, m}} \delta(\omega - \omega_{n, \ell, m}) \right. \\ &\quad \left. + \left[ \tilde{S}_{\ell, |m|}(\mu) - i\zeta_m \tilde{S}_{\ell, -|m|}(\mu) \right] e^{-i\epsilon_{n, \ell, m}} \delta(\omega + \omega_{n, \ell, m}) \right\}; \end{aligned} \quad (2.10)$$

for a finite time string, or for damped oscillations, the delta function is replaced by a line profile of finite width.



### 3 The calculation of the spatial response functions

Instead of using directly equation (2.6) to calculate the response of the observations to a mode of oscillation, it is convenient to use a procedure that follows more closely the actual analysis of the observations. We shall assume throughout that the observations record apparent spectral line shifts, interpreted as velocities, associated with individual pixels in a grid on the solar disk, and that these pixel observations are combined linearly in such a way as to isolate a restricted set of modes. Correspondingly the response calculation can be divided into two parts:

- i) Calculate the responses of the individual pixels to a specific spherical harmonic.
- ii) Combine these in the same way as are the observations.

Clearly step i) includes all the assumptions about the response to the oscillations of the observed spectral line, possibly taking into account the height variation in the eigenfunction, the combined effects of velocities and temperature perturbations on the spectral line, etc., whereas step ii) should be as similar as possible to the analysis of the actual data.

As discussed in Section 2, we assume that the observed spectral line shift is solely due to Doppler shift caused by a purely vertical, height-independent velocity field. Then the response of a pixel  $P_i$  to a single mode is simply the intensity-weighted average of  $v_{0;n,\ell,m}$  (cf. equation (2.4)) over  $P_i$ , and may be written as

$$v_{n,\ell,m}^{(p)}(t;i) = v_{n,\ell,m} \psi_m \left[ \tilde{S}_{\ell,|m|}^{(p)}(i) \cos(\omega_{n,\ell,m} t - \epsilon_{n,\ell,m}) + \zeta_m \tilde{S}_{\ell,-|m|}^{(p)}(i) \sin(\omega_{n,\ell,m} t - \epsilon_{n,\ell,m}) \right] \quad (3.1)$$

where the pixel responses are given by

$$\tilde{S}_{\ell, m}^{(p)}(i) = \frac{(-1)^m}{I^{(p)}(i)} \int_{P_i} I(\theta') c_{\ell, m} P_{\ell}^m(\cos\theta) \cos m\varphi \sin\theta \cos\phi \, dA, \quad \text{for } m \geq 0$$

(3.2)

$$\tilde{S}_{\ell, -m}^{(p)}(i) = \frac{(-1)^m}{I^{(p)}(i)} \int_{P_i} I(\theta') c_{\ell, m} P_{\ell}^m(\cos\theta) \sin m\varphi \sin\theta \cos\phi \, dA, \quad \text{for } m > 0.$$

Here  $I(\theta')$  is the intensity, where  $\theta'$  (given by  $\cos \theta' = \sin\theta \cos\phi$ ) is the angle between the radius vector and line of sight: we use for  $I(\theta')$  the continuum limb darkening function, and as in Christensen-Dalsgaard & Gough approximate it by a quadratic fit (Allen 1973).  $I^{(p)}(i)$  is the intensity associated with  $P_i$ ,

$$I^{(p)}(i) = \int_{P_i} I(\theta') \, dA. \quad (3.3)$$

The combination of the observations is specified by a set of transformation weights  $\{W^{(p)}(i; \mu)\}$ ,  $i = 1, 2, \dots, \mu = 1, 2, \dots$ . Here  $i$  labels the pixels, whereas  $\mu$ , as in Section 2, labels the different transformations. Thus in the  $\mu$ -th combination the observed velocity is

$$\begin{aligned} v_{\text{obs}}^{(t)}(t; \mu) &= \sum_i W^{(p)}(i; \mu) \sum_{n, \ell, m} v_{n, \ell, m}^{(p)}(i; t) \\ &= \sum_{n, \ell, m} V_{n, \ell, m} \psi_m \left[ \tilde{S}_{\ell, |m|}^{(t)}(\mu) \cos(\omega_{n, \ell, m} t - \epsilon_{n, \ell, m}) \right. \\ &\quad \left. + \zeta_m \tilde{S}_{\ell, -|m|}^{(t)}(\mu) \sin(\omega_{n, \ell, m} t - \epsilon_{n, \ell, m}) \right], \end{aligned} \quad (3.4)$$

where the transformed response functions are

$$\tilde{S}_{\ell, m}^{(t)}(\mu) = \sum_i W^{(p)}(i; \mu) \tilde{S}_{\ell, m}^{(p)}(i) \quad (3.5)$$

As in Section 2 we assume that the weights, and hence the transformed responses, are real.

We have computed pixel responses for a set of quadratic pixels with sides  $0.033 \times 0.033 R$ , where  $R$  is the radius of the solar disk. The pixels were aligned in the North-South and East-West directions, with no gaps. All pixels lying entirely within a circle of radius  $0.9 R$  and centred on the disk were included (see Fig. 1). The integrals in equations (3.2) and (3.3) were evaluated using Simpson quadrature in  $\theta$  and  $\phi$ . Further details about the computation are given in Appendix B.

As an example of a transformation we have used a modified version of a scheme originally proposed (but now abandoned) for the analysis of the Fourier Tachometer data (T. M. Brown, private communication). Here the weight is approximately

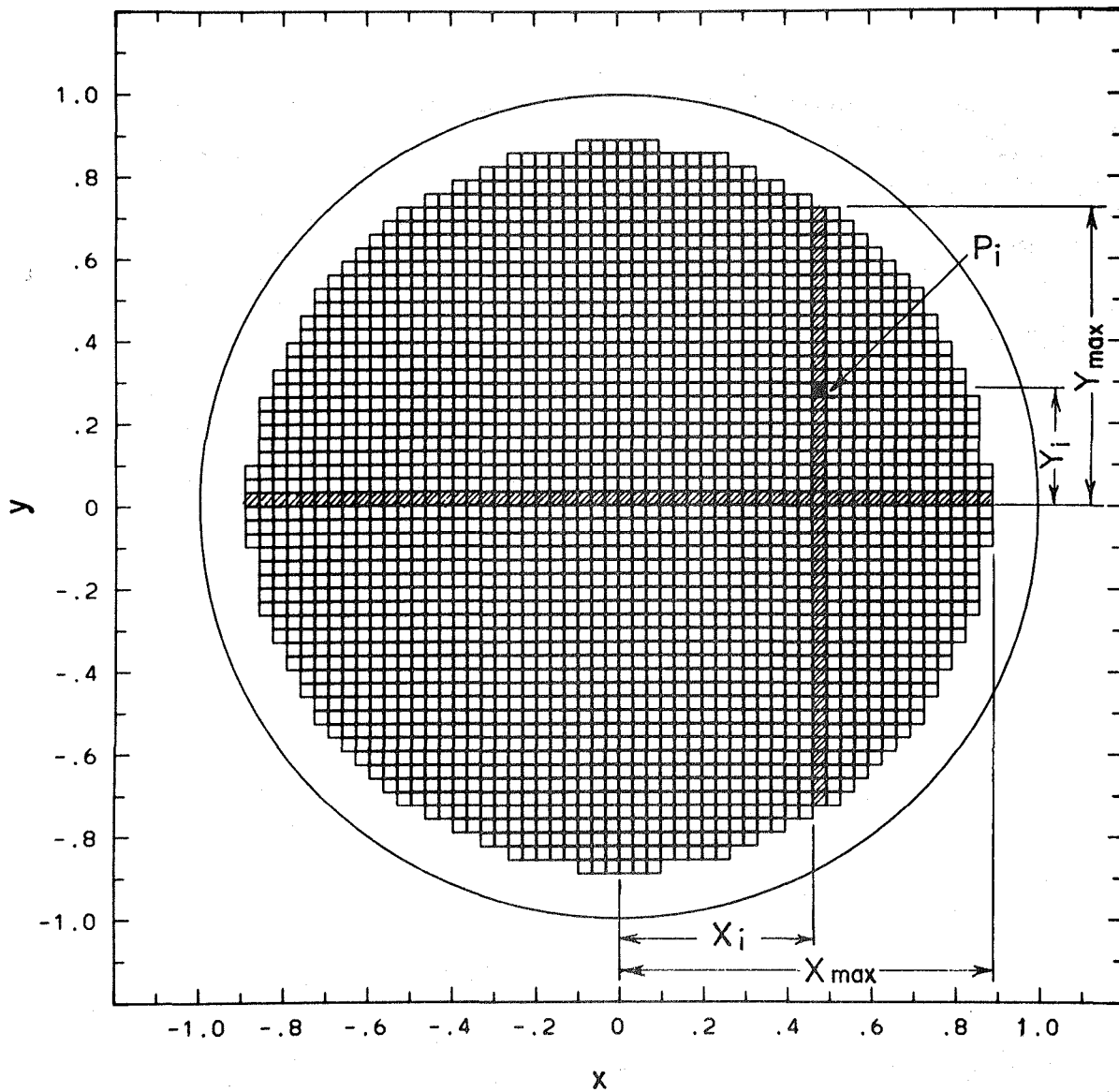
$$W^{(p)}(i; \mu) = \frac{4}{\pi^2} d^2 \left[ (1-\eta_i)^2 (1-\xi_i)^2 \right]^{\frac{1}{2}} U_p(\eta_i) U_q(\xi_i), \quad (3.6)$$

where  $d$  is the size of the pixel, in units of  $R$ :

$$\xi_i = \frac{x_i}{x_{\max}}, \quad \eta_i = \frac{y_i}{y_{\max}}, \quad (3.7)$$

where  $(x_i, y_i)$  are the rectangular coordinates of the centre of the pixel,  $2y_{\max}$  is the height of the pixel column at  $x_i$ , and  $2x_{\max}$  is the total length of the equatorial row of pixels (see Fig. 1).  $U_p$  and  $U_q$  are Chebychev polynomials of the second kind, and the label  $\mu$  corresponds to the pair  $(p, q)$ . This weight function was designed to be somewhat similar to a spherical harmonic, yet with certain features aimed at reducing the noise in the original version of the Fourier Tachometer.

It might be noted that the correspondence between these Chebychev weights and the spherical harmonics is incomplete: the latter have straight horizontal and curved vertical nodal lines, whereas the weight functions given by equation (3.6) have curved horizontal and straight vertical nodal lines. Thus one can expect that these weight functions provide a mode separation inferior to that produced by, for example, weight functions based on spherical harmonics. Therefore they are well suited to test the optimization procedure to be discussed in Section 4.



**Figure 1.** Lay-out of the pixels on the disk of the Sun:  $x = \sin\theta \cos\phi$  and  $y = \cos\theta$  are rectangular coordinates, measured in units of the radius  $R$  of the disk. For the purpose of defining the weight  $W^{(p)}(j; \mu)$  (cf. equation (3.6)) the pixel  $P_i$ , whose centre has coordinates  $(x_i, \mu_i)$ , is shown as black, and the column of pixels to which  $P_i$  belongs, as well as the equatorial row of pixels, are hatched.

When the pixels are placed symmetrically with respect to the centre of the disk, as in the present calculation, the transformed responses satisfy certain selection rules.

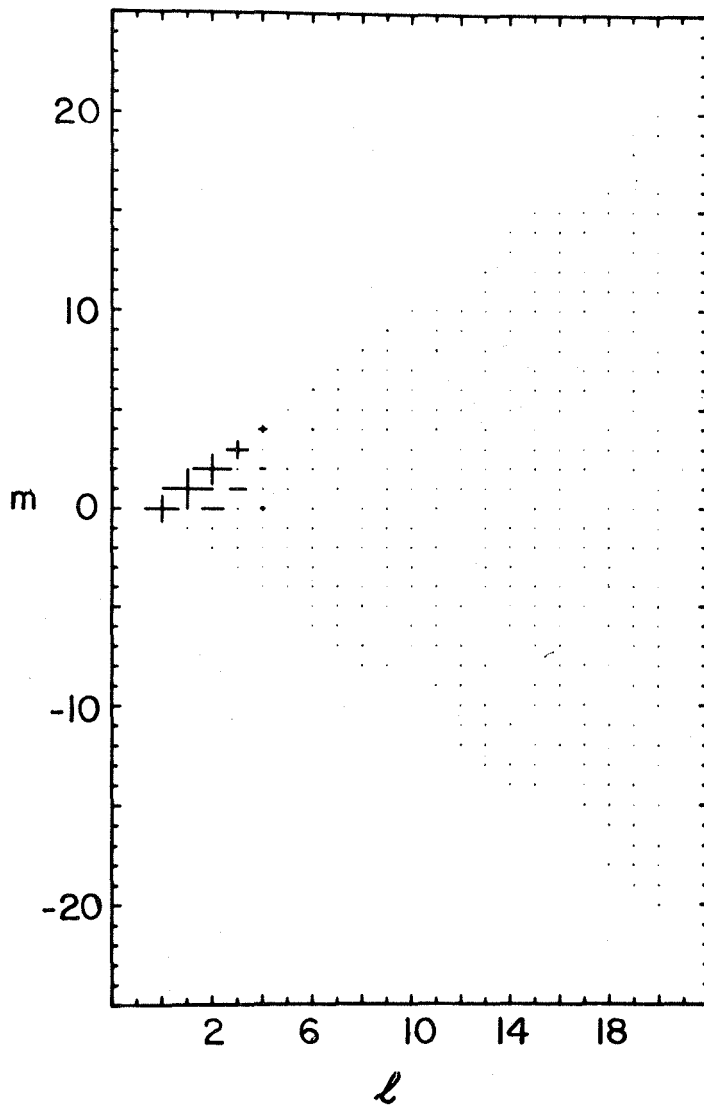


Figure 2. Transformed responses  $\tilde{S}_{l,m}^{(q)}(\mu)$ , for Chebychev weight (cf. equation (3.6)) with  $p = q = 0$ . At each  $l$ ,  $m$  a "+" is shown if  $\tilde{S}_{l,m}^{(q)}(\mu) > 0$ , and a "-" if  $\tilde{S}_{l,m}^{(q)}(\mu) < 0$ . The size of the symbol is proportional to  $|\tilde{S}_{l,m}^{(q)}(\mu)|$ . As  $q$  is even, the responses corresponding to  $m < 0$  are all zero (see Table 1).

due to the symmetry properties of the spherical harmonics. Thus the transformations may be divided into four symmetry classes, such that in each class  $\tilde{S}_{l,m}^{(q)}(\mu)$  is only non-zero for selected  $l$  and  $m$ . These are described in Table 1.

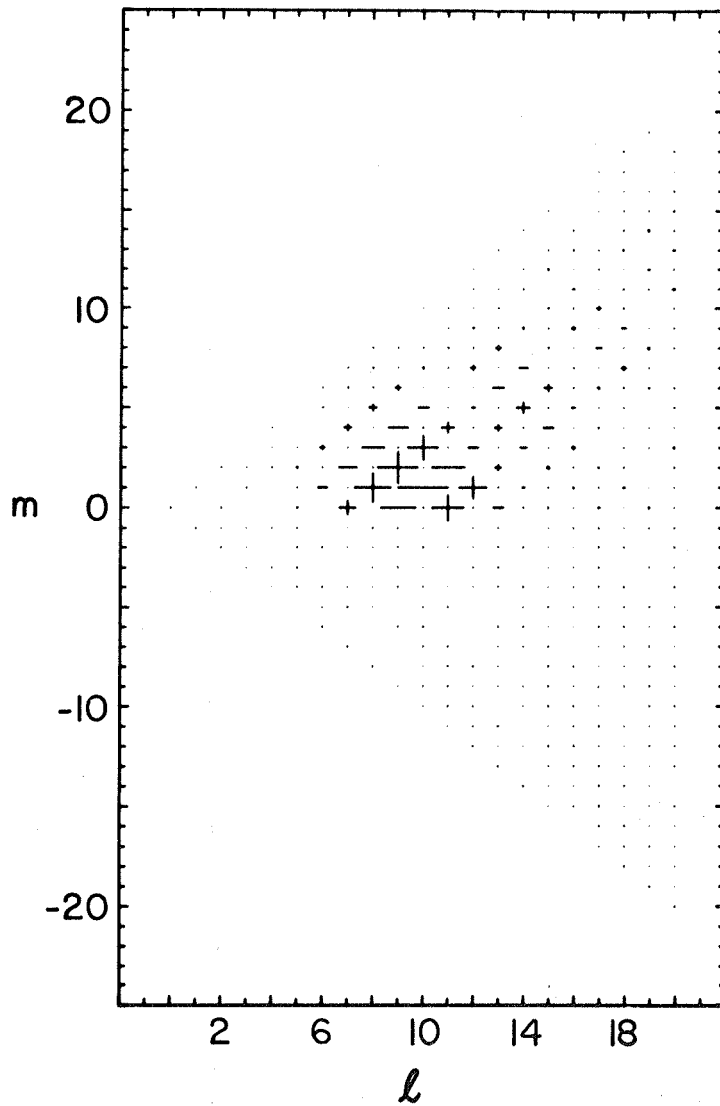


Figure 3. Transformed responses  $\tilde{S}_{l,m}^{(p,q)}(u)$  for Chebychev weight with  $p = 7$ ,  $q = 0$ . See caption to Fig. 2.

Table 1

Symmetry classes for Chebychev transforms

Class	$p, q$ satisfy	$\tilde{S}_{l,m}^{(p,q)}(u) \neq 0$ only for
1	$p$ even, $q$ even	$l-m$ even, $m \geq 0$
2	$p$ odd, $q$ even	$l-m$ odd, $m \geq 0$
3	$p$ even, $q$ odd	$l-m$ even, $m < 0$
4	$p$ odd, $q$ odd	$l-m$ odd, $m < 0$

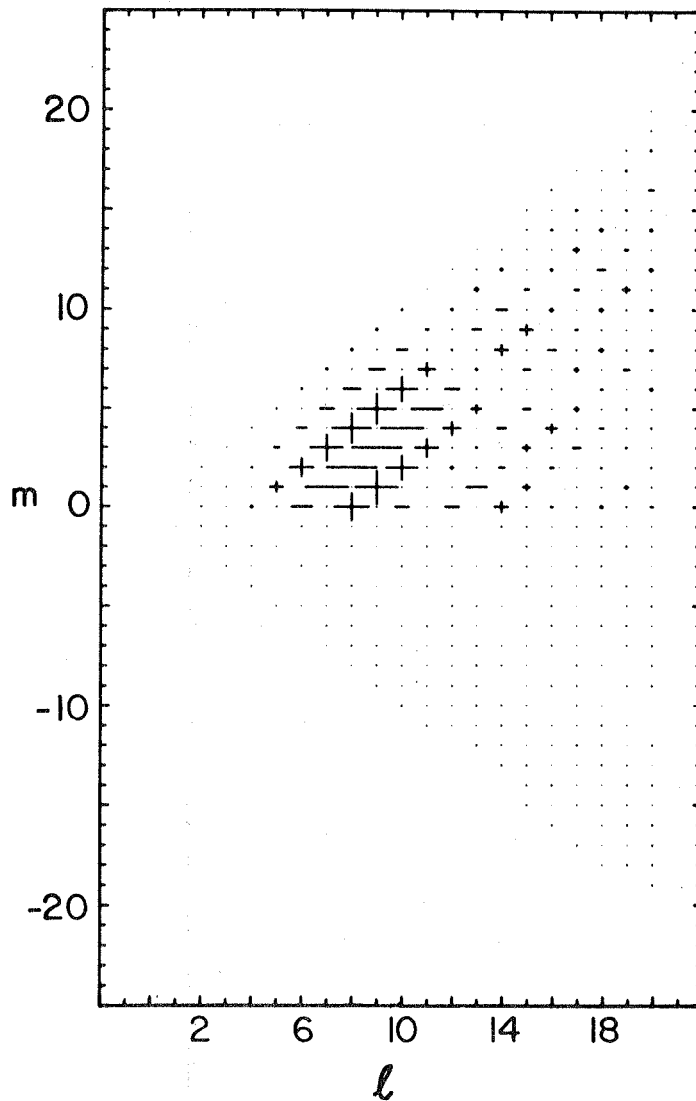
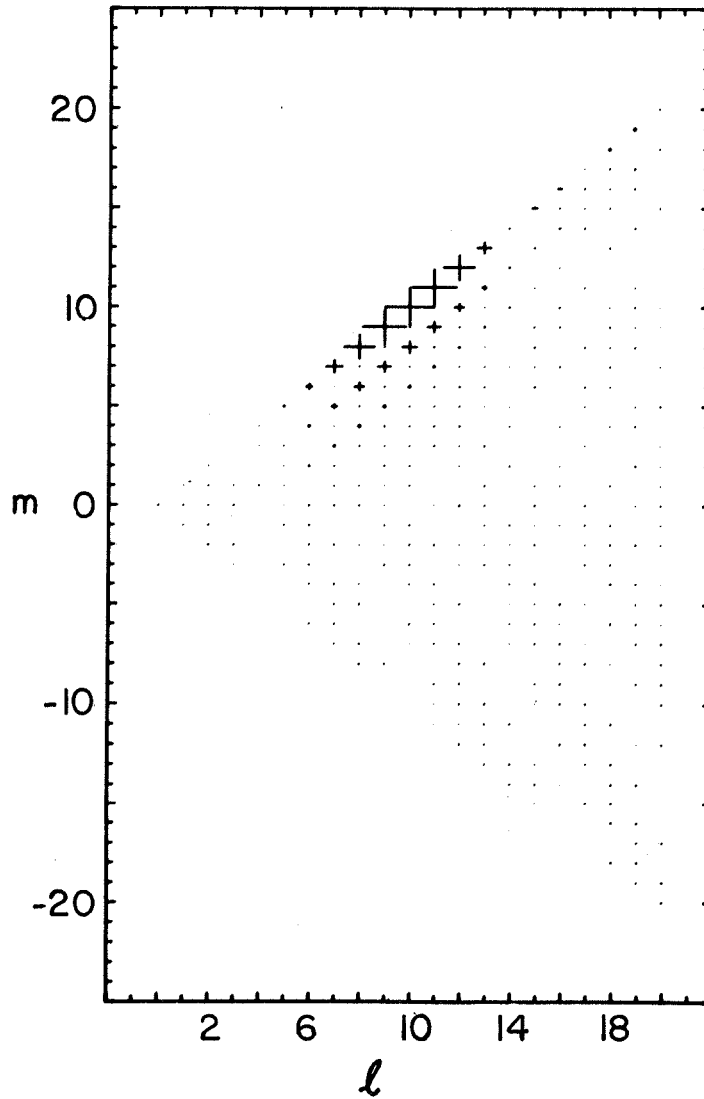


Figure 4. Transformed responses  $\tilde{S}_{l,m}(\mu)$  for Chebychev weight with  $p = 6$ ,  $q = 2$ . See caption to Fig. 2.

Examples of responses transformed with the Chebychev weights are shown on Figs. 2 - 5, for various combinations of  $p$  and  $q$ . It is obvious that these weights succeed in concentrating the sensitivity reasonably well within  $l_c \pm 3$ , where  $l_c$  is the value of  $l$  corresponding to the maximum sensitivity, with a similar range in  $m$ . This might be considered an encouragingly good performance, in view of the qualitative difference in behaviour between the Chebychev weights and the spherical harmonics. However it would still



**Figure 5.** Transformed responses  $\tilde{S}_{l,m}(\mu)$  for Chebychev weight with  $p = 0$ ,  $q = 8$ . See caption to Fig. 2.

cause considerable interference in any one observation between modes separated by 3 in degree, and hence it does not solve the problems with accidental degeneracy discussed in Section 1.



#### 4 Optimal combination of the observations

It may be desirable to combine the results of an initial analysis of the observations, as given in equation (3.4), in such a way as to improve the modal isolation. To do so we seek a combination

$$v_{\text{obs}}^{(o)}(t; \ell_0, m_0) = \sum_{\mu} a_{\mu}(\ell_0, m_0) v_{\text{obs}}^{(t)}(t; \mu) \quad (4.1)$$

that is predominantly sensitive to modes with  $(\ell, m)$  close to  $(\ell_0, m_0)$ . This can be written as in equation (3.4), but with  $\tilde{S}_{\ell, m}^{(t)}(\mu)$  replaced by the optimized response

$$\tilde{S}_{\ell, m}^{(o)}(\ell_0, m_0) = \sum_{\mu} a_{\mu}(\ell_0, m_0) \tilde{S}_{\ell, m}^{(t)}(\ell_0, m_0) \quad (4.2)$$

Similarly, as in equation (2.10), the temporal Fourier transform of  $v_{\text{obs}}^{(o)}$  is

$$\begin{aligned} \tilde{v}_{\text{obs}}^{(o)}(\omega; \ell_0, m_0) &= \frac{1}{2} \sum_{n, \ell, m} V_{n, \ell, m} \psi_m \\ &\times \left\{ \left[ \tilde{S}_{\ell, |m|}^{(o)}(\ell_0, m_0) + i\zeta_m \tilde{S}_{\ell, -|m|}^{(o)}(\ell_0, m_0) \right] e^{i\epsilon_{n, \ell, m} \delta(\omega - \omega_{n, \ell, m})} \right. \\ &\left. + \left[ \tilde{S}_{\ell, |m|}^{(o)}(\ell_0, m_0) - i\zeta_m \tilde{S}_{\ell, -|m|}^{(o)}(\ell_0, m_0) \right] e^{-i\epsilon_{n, \ell, m} \delta(\omega + \omega_{n, \ell, m})} \right\} \end{aligned} \quad (4.3)$$

If the  $a_{\mu}(\ell_0, m_0)$  are assumed to be real, this is also true of the optimized responses  $\tilde{S}_{\ell, m}^{(o)}(\ell_0, m_0)$ . In this case the integrated power associated with the mode  $(n, \ell, m)$  is the same on the positive and negative  $\omega$ -axis and is proportional to

$$P_{n, \ell, m}^{(o)} = V_{n, \ell, m}^2 \left[ S_{\ell, m}^{(o)} \right]^2 \quad (4.4)$$

where

$$\left[ S_{\ell, m}^{(o)} \right]^2 = \left[ \tilde{S}_{\ell, |m|}^{(o)} \right]^2 + \zeta_m^2 \left[ \tilde{S}_{\ell, -|m|}^{(o)} \right]^2, \quad (4.5)$$

and we may take  $S_{\ell, m}^{(o)} \geq 0$ . It should also be noticed that

$$S_{\ell, m}^{(o)} = S_{\ell, -m}^{(o)}; \quad (4.6)$$

thus, because of the purely real weighting, there is no distinction between modes with positive and negative  $m$ . On the other hand, by using complex  $a_n(\ell_0, m_0)$  it is possible predominantly to have sensitivity to positive  $m$  on the positive  $\omega$ -axis; the negative  $\omega$ -axis is then predominantly sensitive to negative  $m$ . Such complex weighting is also implicit in the Fourier transform in  $\phi$  employed by, for example, Rhodes et al (1977). Complex weighting may be employed in the optimization procedure; however we shall not consider it further here.

To determine the coefficients  $a_n(\ell_0, m_0)$  we use the Backus & Gilbert (1970) technique, as did Gough & Latour (1984). Specifically we determine the  $a_n(\ell_0, m_0)$  that minimize

$$\mathcal{V} = \sum_{\ell, m} U(\ell, |m|; \ell_0, m_0) \left[ S_{\ell, m}^{(o)}(\ell_0, m_0) \right]^2 \quad (4.7)$$

subject to

$$\mathcal{J} \equiv \sum_{\ell, m} \left[ S_{\ell, m}^{(o)}(\ell_0, m_0) \right]^2 = 1. \quad (4.8)$$

Here  $U(\ell, |m|; \ell_0, m_0)$  is a suitable weight function that is "small" at  $\ell = \ell_0$ ,  $|m| = |m_0|$ , and "large" elsewhere. Notice that in accordance with the symmetry in equation (4.6) we apply equal weight to  $m$  and  $-m$ . The result of the minimization of  $\mathcal{V}$  subject to  $\mathcal{J} = 1$  is to concentrate as far as possible  $S_{\ell, m}^{(o)}$  at  $\ell = \ell_0$  and  $|m| = m_0$ ; the choice of the actual value of the normalization in equation (4.8) is clearly arbitrary.

This procedure is not strictly equivalent to the normal Backus & Gilbert technique (see the review by Gough, these proceedings). In the latter the normalization is applied directly to the linear combination of, for example, the rotational kernels; in contrast we have normalized on the sum of squares of the combined responses, thereby, in a

loose sense, fixing the total power. This modification is natural, in view of the fact that both positive and negative responses are involved: a normalization that included the sign of the optimized responses would have little meaning.

The optimized responses are to some extent determined by the choice of  $U$ . We have used

$$U(\ell, |m|; \ell_0, m_0) = \frac{X}{1+X}, \quad (4.9a)$$

where

$$X = \frac{1}{2} \left[ (\ell - \ell_0)^2 + (|m| - |m_0|)^2 \right]; \quad (4.9b)$$

this has the desirable property that equal weight is applied to all  $(\ell, m)$  pairs far from  $(\ell_0, m_0)$ . However no extensive experimentation has been carried out so far to determine the optimum form of  $U$ ; possibly different forms may be indicated in different situations, to obtain specific properties of the resulting response.

It is easy to show that the minimization of equation (4.7) subject to equation (4.8) is equivalent to determining the smallest eigenvalue  $\lambda_{min}$  of the generalized eigenvalue problem

$$\sum_{\nu} A_{\mu, \nu} a_{\nu} = \lambda \sum_{\nu} Q_{\mu, \nu} a_{\nu} \quad (4.10)$$

where we have suppressed the arguments  $(\ell_0, m_0)$  of  $a_{\nu}$ ; here

$$A_{\mu, \nu} = \sum_{\ell, m} U(\ell, |m|; \ell_0, m_0) \tilde{S}_{\ell, m}^{(t)}(\mu) \tilde{S}_{\ell, m}^{(t)}(\nu), \quad (4.11)$$

and

$$Q_{\mu, \nu} = \sum_{\ell, m} \tilde{S}_{\ell, m}^{(t)}(\mu) \tilde{S}_{\ell, m}^{(t)}(\nu). \quad (4.12)$$

Then the minimum of  $\mathcal{V}$  is  $\mathcal{V}_{min} = \lambda_{min}$ . If the spatial responses  $\tilde{S}_{\ell, m}^{(t)}(\mu)$  are linearly independent, the matrix  $\{Q_{\mu, \nu}\}$  is positive definite, and equation (4.10) may be solved using standard techniques. The generalized eigenvalue problem arises because of the quadratic normalization used here; in ordinary Backus & Gilbert inversion equation (4.10)

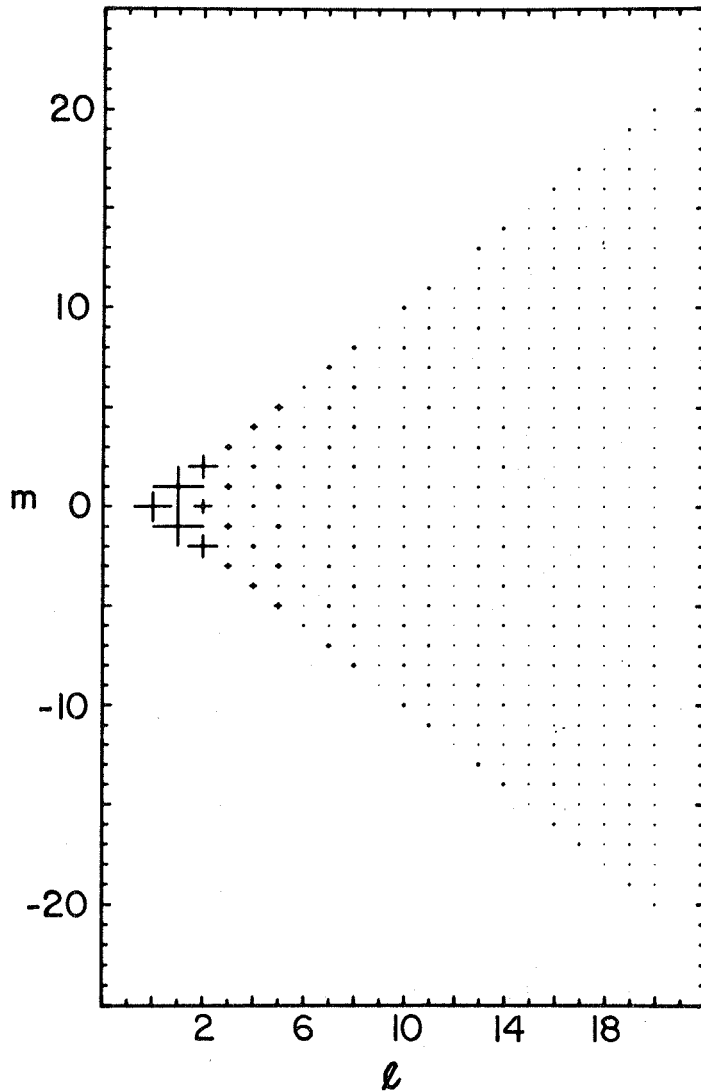
is replaced by a set of linear equations for the coefficients  $\{a_\nu\}$ .

It may be shown that when the input responses separate into symmetry classes, as in the case of the Chebychev weighted responses (cf. Table 1), this is true also for the optimized responses: thus only coefficients  $a_\nu$  for which  $\nu$  belongs to a particular symmetry class are non-zero. It is then natural to seek the minimum of  $\mathcal{V}$  in the symmetry class corresponding to the  $(\ell_0, m_0)$  desired. The symmetry may be exploited computationally by condensing the matrices  $\{A_{\mu,\nu}\}$  and  $\{Q_{\mu,\nu}\}$  to include only elements for which  $\mu$  and  $\nu$  have the correct symmetry.

We have used the optimization technique on the Chebychev-weighted responses computed in Section 3. Chebychev indices  $p$  and  $q$  (cf. equation (3.6)) between 0 and 9 were used, and the summation over  $\ell$  and  $m$  included values between 0 and 50. A few typical optimized amplitude responses  $S_{\ell_0, m_0}^{(p, q)}(\ell_0, m_0)$  are shown on Figs. 6 - 9. By comparing with Figs. 2 - 5 it is obvious that a substantial improvement in concentration has been obtained. In particular the sensitivity is now quite low at  $\ell_0 \pm 3$ , and so the interference due to frequency degeneracy between modes separated by 3 in  $\ell$  should be substantially reduced. In fact in the "hide-and-seek" experiment by Mihalas et al (these proceedings), which used these responses, the modes could be uniquely identified and their frequencies were recovered within the frequency resolution given by the length of the time series: this indicates that interference had no serious effects.

These results show that well-concentrated responses can be obtained on the basis of the Chebychev-weighted responses. However the resulting coefficients  $a_\mu(\ell_0, m_0)$  corresponded to an amplification by a factor of up to about two of the errors contained in the original Chebychev-weighted data. Similar problems with error amplification occur in ordinary Backus & Gilbert inversion (see Christensen-Dalsgaard & Gough, these proceedings). As there, they may be solved by adding to  $\mathcal{V}$  in equation (4.7) a measure of the variance, with a weight that can be adjusted to get a reasonable trade-off between the degree of concentration and the error amplification. This was done by Gough & Latour (1984) for the optimized combination of limb observations. We plan to pursue it in future.

It might be noted that although the  $a_\mu$  are fairly large, the weights applied to the individual pixels are reasonable. From equations (4.2) and (3.5) the optimized responses may be written as



**Figure 6.** Optimized responses  $S_{l,m}^{(o)}(\ell_0, m_0)$ , defined in equations (4.4) and (4.5), for the amplitude deduced from a power spectrum, with  $\ell_0 = m_0 = 0$ . The responses are shown in the same way as on Fig. 2 - 5 (see caption to Fig. 2). It should be noticed, however, that the  $S_{l,m}^{(o)}$  are defined such as to be non-negative and independent of the sign of  $m$  (cf. equation (4.6)).

$$\tilde{S}_{\ell,m}^{(o)}(\ell_0, m_0) = \sum_i W^{(o)}(i; \ell_0, m) \tilde{S}_{\ell,m}^{(p)}(i), \quad (4.13)$$

where

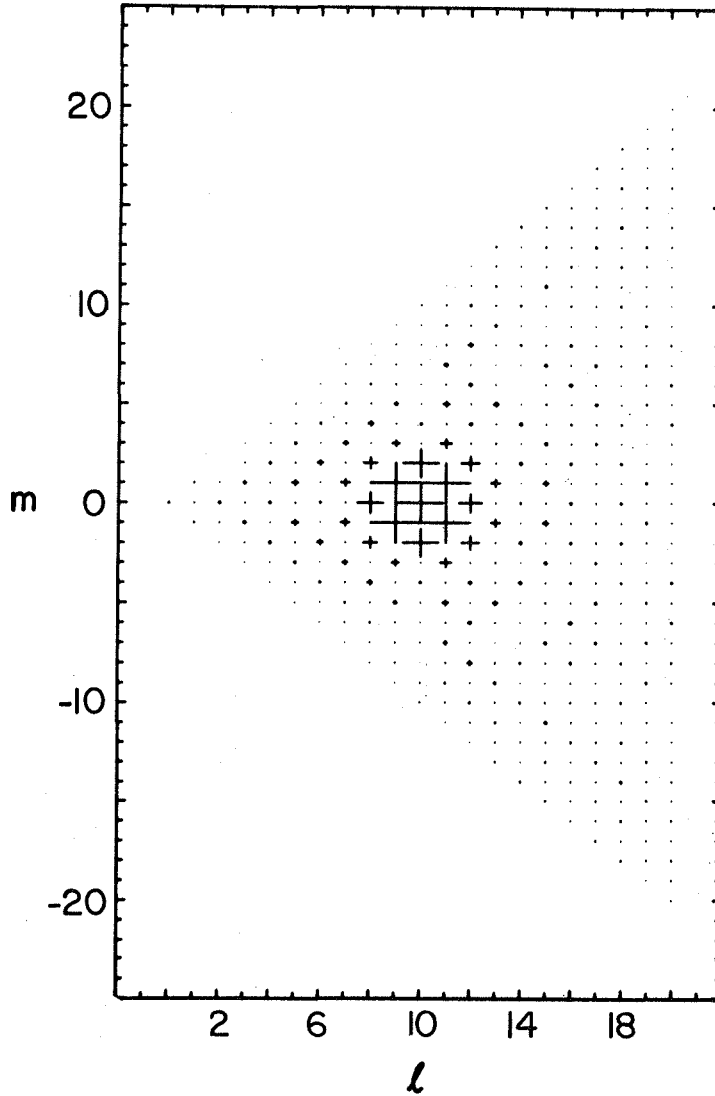


Figure 7. Optimized responses  $S_{i;l_0,m_0}^{(o)}$  with  $l_0 = 10$ ,  $m_0 = 0$ . See caption to Fig. 6.

$$W^{(o)}(i;l_0,m_0) = \sum_{\mu} a_{\mu}(l_0,m_0) W^{(p)}(i;\mu), \quad (4.14)$$

and here the weights  $W^{(o)}(i;l_0,m_0)$  are comparable in size to the original Chebychev weights  $W^{(p)}(i;\mu)$ . Thus it is possible to combine the pixel data to obtain a fairly high degree of concentration, without serious error amplification. As might have been expected, the optimized weights are, at least qualitatively, similar to spherical harmonics.

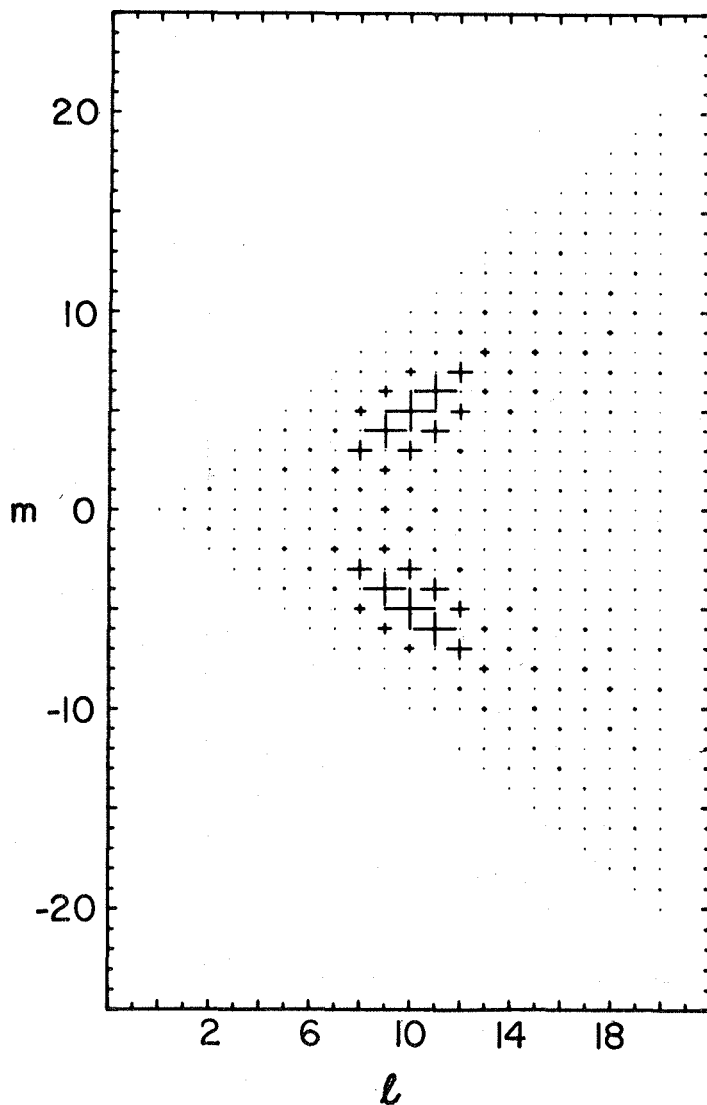


Figure 8. Optimized responses  $S_{i,j}^{*}(\ell_0, m_0)$  with  $\ell_0 = 10$ ,  $m_0 = 5$ . See caption to Fig. 6.

### 5 Optimization of bidirectional responses

It is straightforward to extend the techniques discussed in Section 4 to the combination of two different sets of observations. Thus we can investigate the improvement that would result by having, as in the original concept for the

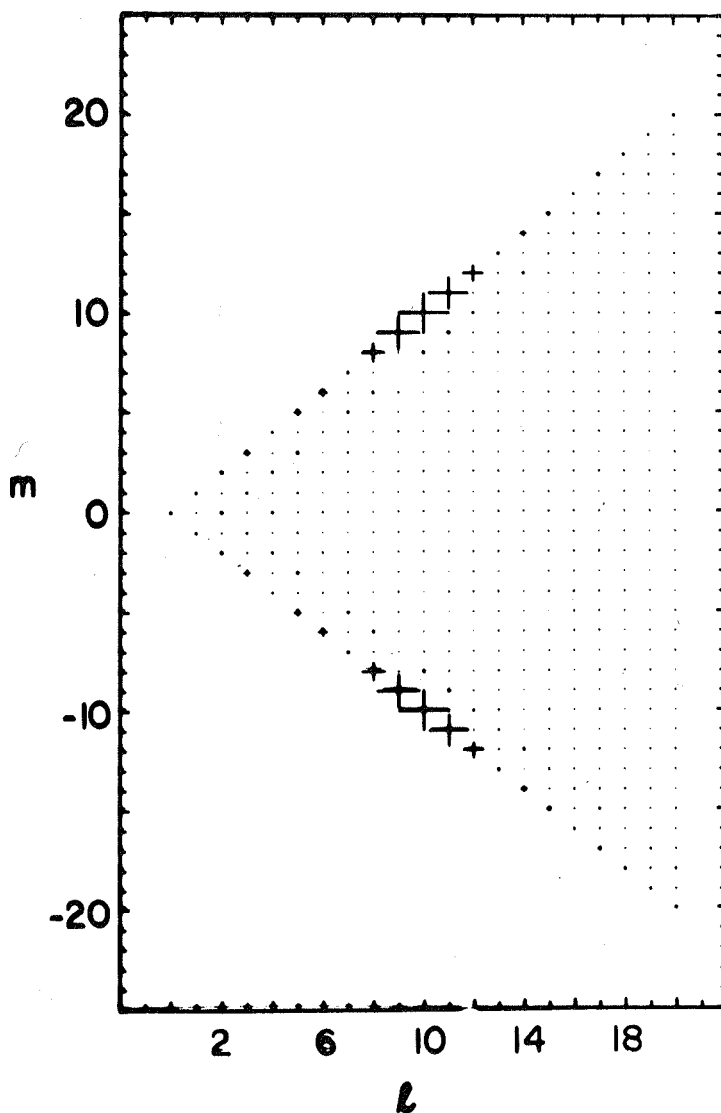


Figure 9. Optimized responses  $S_{l,m}^{(1)}(l_0, m_0)$  with  $l_0 = 10$ ,  $m_0 = 10$ . See caption to Fig. 6.

SDO, observations from two separate directions relative to the Sun.

For simplicity we assume that both observing stations lie in the plane of the solar equator, the angle between them being  $\Phi$ ; we also assume the observational set-up to be identical for the two stations. It is then easy to show that the Fourier transform of the time string obtained at the second station is



$$\begin{aligned}
\tilde{u}_{\text{obs}}^{(t)}(\omega; \mu) &= \frac{1}{2} \sum_{n, \ell, m} V_{n, \ell, m} \psi_m \\
&\times \left\{ \left[ \tilde{S}_{\ell, |m|}^{(t)}(\mu) + i\zeta_m \tilde{S}_{\ell, -|m|}^{(t)}(\mu) \right] e^{i\epsilon_{n, \ell, m}} e^{im\phi} \delta(\omega - \omega_{n, \ell, m}) \right. \\
&+ \left. \left[ \tilde{S}_{\ell, |m|}^{(t)}(\mu) - i\zeta_m \tilde{S}_{\ell, -|m|}^{(t)}(\mu) \right] e^{-i\epsilon_{n, \ell, m}} e^{-im\phi} \delta(\omega + \omega_{n, \ell, m}) \right\}; \quad (5.1)
\end{aligned}$$

the Fourier transform  $\tilde{v}_{\text{obs}}^{(t)}(\omega; \mu)$  of the time string at the first station is given by an equation corresponding to equation (2.10). We now seek a combination

$$\begin{aligned}
\tilde{v}_{\text{obs}}^{(o)}(\omega; \ell_0, m_0) &= \sum_{\mu} a_{\mu}(\ell_0, m_0) \tilde{v}_{\text{obs}}^{(t)}(\omega; \mu) \\
&+ \sum_{\mu'} b_{\mu'}(\ell_0, m_0) \tilde{u}_{\text{obs}}^{(t)}(\omega; \mu') \quad (5.2)
\end{aligned}$$

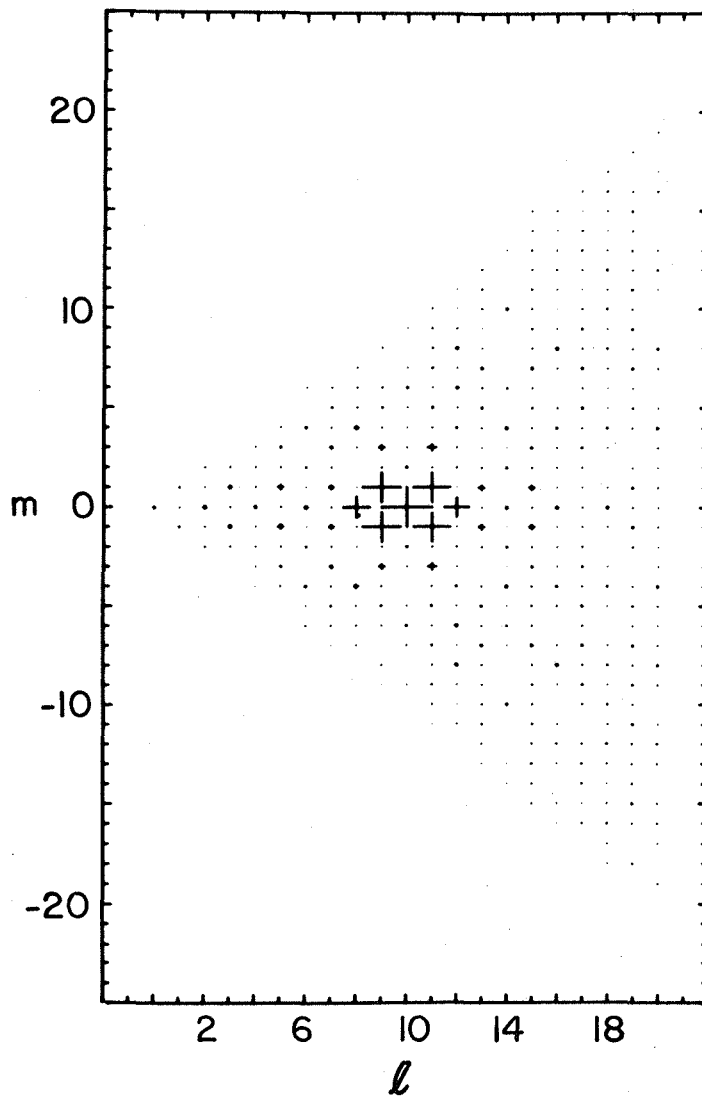
that optimizes the concentration: as in Section 4 we assume that all coefficients are real. This may be written as in equation (4.3), but with

$$\begin{aligned}
\tilde{S}_{\ell, m}^{(o)}(\ell_0, m_0) &= \sum_{\mu} a_{\mu}(\ell_0, m_0) \tilde{S}_{\ell, m}^{(t)}(\mu) \\
&+ \sum_{\mu'} b_{\mu'}(\ell_0, m_0) \hat{S}_{\ell, m}^{(t)}(\mu'), \quad (5.3)
\end{aligned}$$

where

$$\hat{S}_{\ell, m}^{(t)}(\mu) = \cos m\phi \tilde{S}_{\ell, m}^{(t)}(\mu) - \sin m\phi \tilde{S}_{\ell, -m}^{(t)}(\mu). \quad (5.4)$$

As before the coefficients  $\{a_{\mu}(\ell_0, m_0)\}$ ,  $\{b_{\mu'}(\ell_0, m_0)\}$  are determined by minimizing equation (4.7) subject to equation (4.8). This again leads to a generalized eigenvalue problem, where the eigenvector is now the combination  $\{a_{\mu};$



**Figure 10.** Bidirectional optimized responses  $S_{\ell_0, m_0}^{(2)}$  for the amplitude deduced from a power spectrum, with  $\ell_0 = 10$ ,  $m_0 = 0$ , obtained by combining data from two stations in the solar equatorial plane separated by  $\Phi = 90^\circ$ . The responses are shown in the same way as on Figs. 7 - 10.

$\Delta_{\ell}$ ). It should be noticed that the symmetry conditions in  $q$  (or equivalently  $m$ ) no longer apply, because of the combination of positive and negative  $m$  in the calculation of  $\hat{S}_{\ell, m}^{(2)}(\mu)$ . However there is still a separation into two symmetry classes according to the parity of  $\ell - m$ .

For the Chebychev-weighted responses it was found computationally that the coefficients satisfied symmetry relations, which may be written as

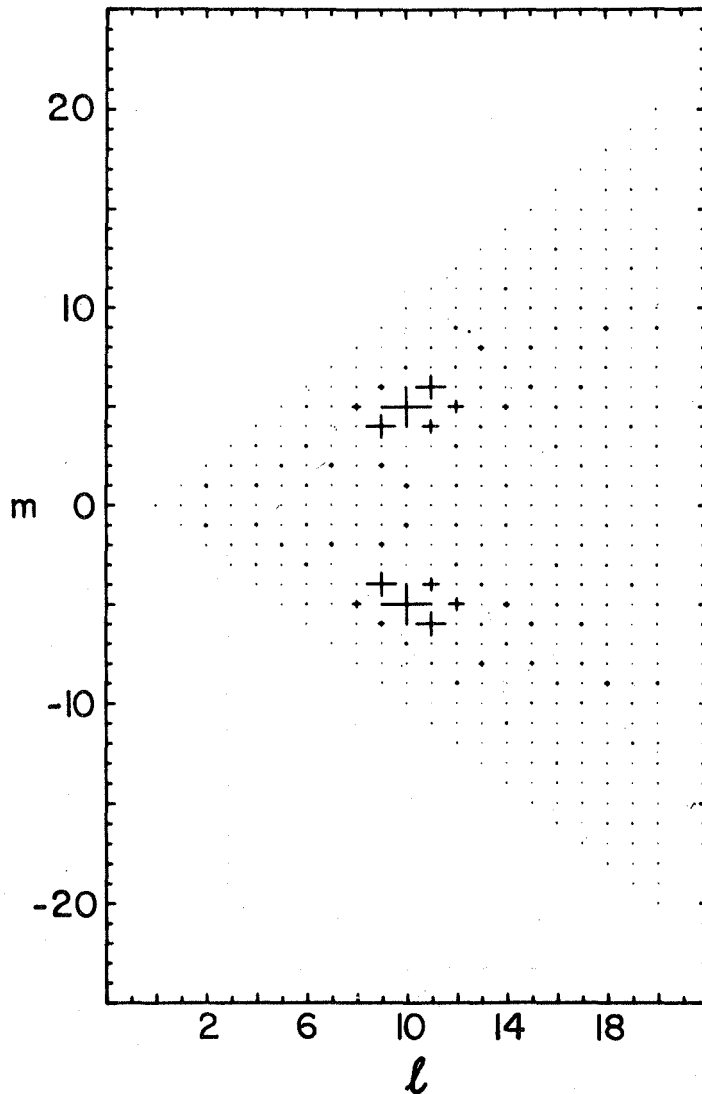


Figure 11. Bidirectional optimized responses  $S_{\ell, m}^{(q)}(\ell_0, m_0)$ , with  $\ell_0 = 10$ ,  $m_0 = 5$ . See caption to Fig. 10.

$$b_{\mu} = \sigma(\ell, m) (-1)^q a_{\mu} ; \quad (5.5)$$

in the relatively few cases considered  $\sigma$  was  $-1$  only for even  $\ell$  and odd  $m$ , and  $1$  otherwise. That relations of this nature should exist seems reasonable, considering the symmetry of the problem: in particular it is natural that the symmetry in  $\phi$  of the weight, given by the value of  $q$ ,

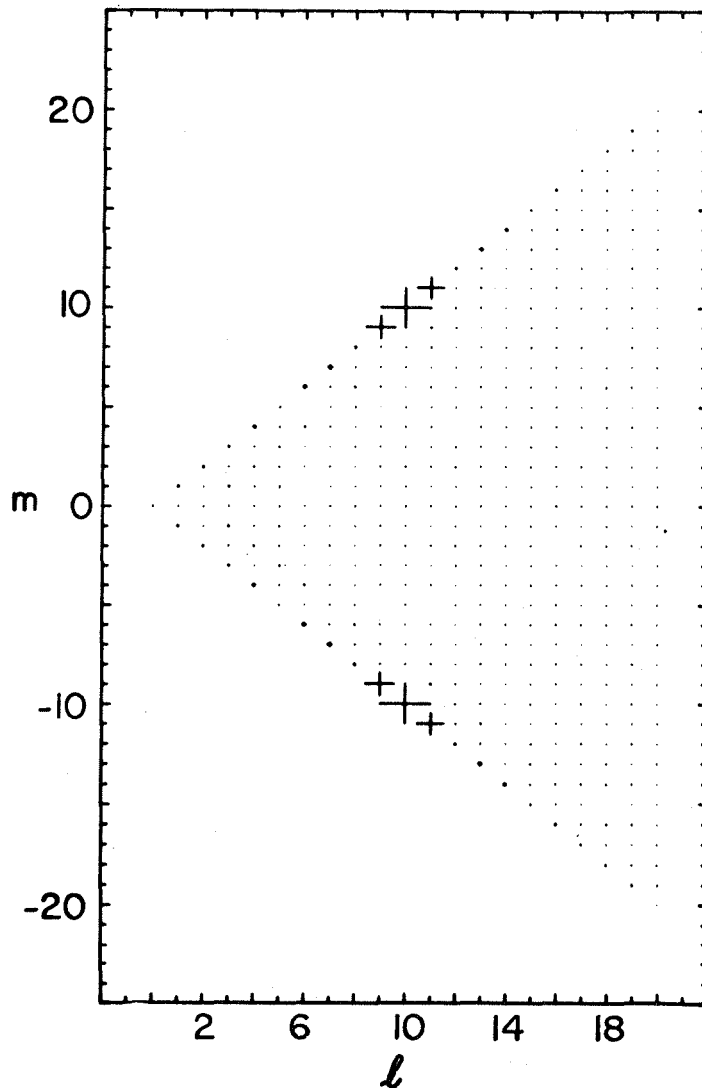


Figure 12. Bidirectional optimized responses  $S_{i,j}^m(\ell_0, m_0)$ , with  $\ell_0 = 10$ ,  $m_0 = 10$ . See caption to Fig. 10.

should come in. However we have not attempted to prove that these relations are always true, nor have we made an exhaustive numerical search to try to verify them for a representative set of cases. Such relations, when firmly established, can be used to reduce the effort required to calculate the optimum combinations. It is obvious that this can apply only when the two observing stations are identical.

Examples of combined bidirectional, Chebychev-weighted

responses are shown on Figs. 10 - 12. These all assumed  $\Phi = 90^\circ$ ; results for  $\Phi = 60^\circ$  and  $120^\circ$  have also been computed, and are similar. The improvement over the mono-directional responses shown on Figs. 7 - 9 is quite striking. It should be noticed, however, that because the extension of the baseline is in  $\phi$  this improvement occurs principally in  $m$ . Thus for  $m_0 = 0$  the sensitivity is as high for  $\ell = \ell_0 \pm 2$ .  $m = 0$ , as in the mono-directional case. In contrast, for  $\ell_0 = m_0$  the responses are essentially confined to  $\ell = m$ , and so the reduction of the spread in  $m$  of the response causes a reduction of the spread in  $\ell$  to  $\ell = \ell_0, \ell_0 \pm 1$ . Further improvement in the mode isolation in  $\ell$  would probably require additional observing stations away from the plane of the solar equator, but this seems entirely impractical in the foreseeable future.

## 6 Discussion and conclusion

We have attempted a reasonably realistic simulation of the observing process, in the sense that the spatial responses are based on averages of the spherical harmonics over individual pixels. This evidently neglects the effects of the variation with height of the eigenfunction, as well as the details of the perturbation in the spectral line profile. Such details are undoubtedly important in obtaining reliable measurements of the oscillation amplitudes (expressed in terms of, e.g., photospheric velocities) from the observed apparent line shifts; but they are unlikely to change qualitatively the conclusions obtained here.

Our main result is that one can combine velocity observations made on a set of pixels lying within 0.9 solar radii from the centre of the solar disk, in such a way as to concentrate the sensitivity within about 5 consecutive values of  $\ell$ . Thus frequencies of modes at the central degree  $\ell_c$  are not significantly contaminated by modes with almost coincident frequencies at  $\ell = \ell_c \pm 3$ . This seems to weaken considerably the argument that bidirectional observations are required to resolve adequately the modes of degree 5 to 15. In fact the simulated data analysis by Mihalas et al (these proceedings), which used the spatial responses obtained here, showed that full identification of the modes and determination of their frequencies to within the expected accuracy were possible. Furthermore the observations of Duvall & Harvey (these proceedings) have yielded individual frequencies for nearly zonal modes of degree up to 14.

To achieve such a high degree of modal isolation it is important to include as large an area of the solar disk as possible. By decreasing the observed area the widths of the

responses are correspondingly increased; thus if only pixels within  $0.8 R$  are used, the responses at  $\ell_c \pm 3$  are about 40 per cent of the response at  $\ell_c$ , and this would cause fairly serious interference in the power spectra. The size of the observing area may have to be restricted because of supergranular noise (Ulrich, these proceedings). Thus it is important that a careful analysis of such effects, based on simulations or, ideally, observations be carried out.

The responses were calculated as optimized combinations of Chebychev-weighted pixel averages. These combinations were found by using the Backus & Gilbert technique of optimal averaging (see also Gough & Latour 1984). The Chebychev weighting has been proposed by T. M. Brown to compensate for certain features of a specific observing scheme, and should probably not be generally applied. A more natural choice would be to analyze the observations with spherical harmonics; in fact the optimal combinations of the Chebychev weights were qualitatively similar to spherical harmonics. However it is not obvious that spherical harmonic weighting over a restricted portion of the solar surface provides responses with the best possible modal isolation; the Backus & Gilbert technique may be employed to seek optimal combinations of such responses. In any case the results obtained here indicate the potential of this technique for constructing optimum analysis schemes for solar oscillation observations. A natural use of the technique would be to determine directly the weight to be applied to individual pixels in observations with limited spatial resolution. In principle we could have done so here, but due to the large number of pixels involved the computational effort would have been excessive.

We have also considered the combination of observations made from two different directions. The resulting responses showed a considerable improvement in the mode isolation in  $m$ , and as a result there was also some improvement in the isolation in  $\ell$ , especially for  $m \approx \ell$ . Essentially no improvement in  $\ell$  was found for  $m \approx 0$ , where the sensitivity was still considerable at  $\ell = \ell_c \pm 2$ . Indeed it is entirely reasonable that an extension of the base-line of the observations in the  $\phi$ -direction should have little effect on the separation of zonal harmonics. Thus in fact bi-directional viewing only partially meets the goal of improving the modal isolation.

No attempts have been made to restrict the size of the coefficients in the optimized combinations by using the trade-off technique (e.g. Gough, these proceedings). As a result there was some amplification of the errors in the original Chebychev-weighted data; on the other hand the combined weights applied to the individual pixels were

reasonably small, and so no amplification of the errors associated with the pixel data should occur. Trade-off can easily be implemented in the optimization (cf. Gough & Latour 1984), and we intend to consider this in future.

**Acknowledgements.** I am very grateful to D. O. Gough for proposing the use of the Backus & Gilbert technique to optimize the observations, and to T. M. Brown for many discussions on the intricacies of data analysis. T. M. Brown and B. W. Mihalas are thanked for reading and commenting on an earlier version of the paper.

## References

- Abramowitz, M. & Stegun, I. A., 1965. Handbook of Mathematical Functions (Dover Publications, Inc., New York).
- Allen, C. W., 1973. Astrophysical Quantities (Athlone Press, London).
- Backus, G. & Gilbert, F., 1970. Phil. Trans., **A266**, 123.
- Bos, R. J. & Hill, H. A., 1983. Solar Phys., **82**, 89.
- Brookes, J. R., Isaak, G. R. & van der Raay, H. B., 1978. Mon. Not. R. astr. Soc., **185**, 19.
- Brown, T. M., Stebbins, R. T. & Hill, H. A., 1978. Astrophys. J., **223**, 324.
- Christensen-Dalsgaard, 1982. Proc. Conference on Pulsations in Classical and Cataclysmic Variable Stars, p. 99 (ed. J. P. Cox and C. J. Hansen, JILA, Boulder CO).
- Christensen-Dalsgaard, J. & Gough, D. O., 1982. Mon. Not. R. astr. Soc., **198**, 141.
- Deubner, F.-L., Ulrich, R. K. & Rhodes Jr., E. J., 1979. Astron. Astrophys., **72**, 177.
- Dicke, R. H., 1982. Nature, **300**, 693.
- Duvall, T. L. & Harvey, J. W., 1983. Nature, **302**, 24.
- Dziembowski, W., 1977. Acta Astron., **27**, 203.
- Edmonds, A. R., 1960. Angular Momentum in Quantum Mechanics, 2. Ed. (Princeton University Press, Princeton N. J.).
- Gough, D. O., 1980. Lecture Notes in Physics, **125**, 277 (ed. H. A. Hill and W. A. Dziembowski, Springer-Verlag, Berlin and Heidelberg).
- Gough, D. O., 1982. Nature, **298**, 350.
- Gough, D. O. & Latour, J., 1984. Solar Phys., in press.
- Hill, F., Toomre, J. & November, L. J., 1983. Solar Phys., **82**, 411.
- Hill, H. A., 1978. The New Solar Physics, Chapter 5 (ed. J. A. Eddy, Westview, Boulder CO).

- Isaak, G. R., 1982. Nature. **296**, 130.  
 Loumos, G. L. & Deeming, T. J., 1978. Astrophys. Space Sci., **56**, 285.  
 Rhodes Jr., E. J., Ulrich, R. K. & Simon, G. W., 1977. Astrophys. J., **218**, 901.  
 Yerle, R., 1981.. Astron. Astrophys., **100**, L23.

## Appendix A. Rotation of the spatial response functions

As discussed in Section 2 the modes of oscillation should be described in terms of spherical harmonics in a coordinate system whose polar axis is aligned with the rotation axis; we label this system  $\mathcal{A}'$ . On the other hand it is computationally very convenient to evaluate the spatial responses in a system  $\mathcal{A}$  with polar axis in the plane of the sky. The transformation between the two systems can be carried out using rotation matrices for spherical harmonics (Edmonds 1960; see also Hill 1978).

The coordinate transformation taking coordinates in  $\mathcal{A}'$  into coordinates in  $\mathcal{A}$  can be described in terms of a set  $(\alpha, \beta, \gamma)$  of Euler angles (cf. Edmonds 1960, Chapter 1, for the definition). Then equation (2.1) for the velocity field associated with a single mode is valid in the system  $\mathcal{A}'$ , and instead of equation (2.7) we obtain

$$v_{\text{obs}}(t; \mu) = \sum_{n, \ell, m} V_{n, \ell, m} \psi_m \left[ \tilde{S}'_{\ell, |m|}(\mu) \cos(\omega_{n, \ell, m} t - \epsilon_{n, \ell, m}) + \zeta_m \tilde{S}'_{\ell, -|m|}(\mu) \sin(\omega_{n, \ell, m} t - \epsilon_{n, \ell, m}) \right], \quad (\text{A.1})$$

where

$$\tilde{S}'_{\ell, m}(\mu) = \sum_{m'} \tilde{\mathcal{D}}_{m' m}^{(\ell)}(\alpha, \beta, \gamma) \tilde{S}_{\ell, m'}(\mu), \quad (\text{A.2})$$

and the  $\tilde{S}_{\ell, m}(\mu)$  are still evaluated using equations (2.8). Here  $\tilde{\mathcal{D}}_{m' m}^{(\ell)}(\alpha, \beta, \gamma)$  is a real rotation matrix which may be derived from the complex rotation matrix  $\mathcal{D}_{m' m}^{(\ell)}(\alpha, \beta, \gamma)$  used by Edmonds (1960).

We consider the special case where the transformation corrects only for the inclination of the rotation axis, corresponding to a rotation around the intersection between the solar equatorial plane and the plane of the sky. Here  $\alpha = \gamma = 0$ , and  $\beta = -\theta_r$ , where  $\theta_r$  is the inclination of the



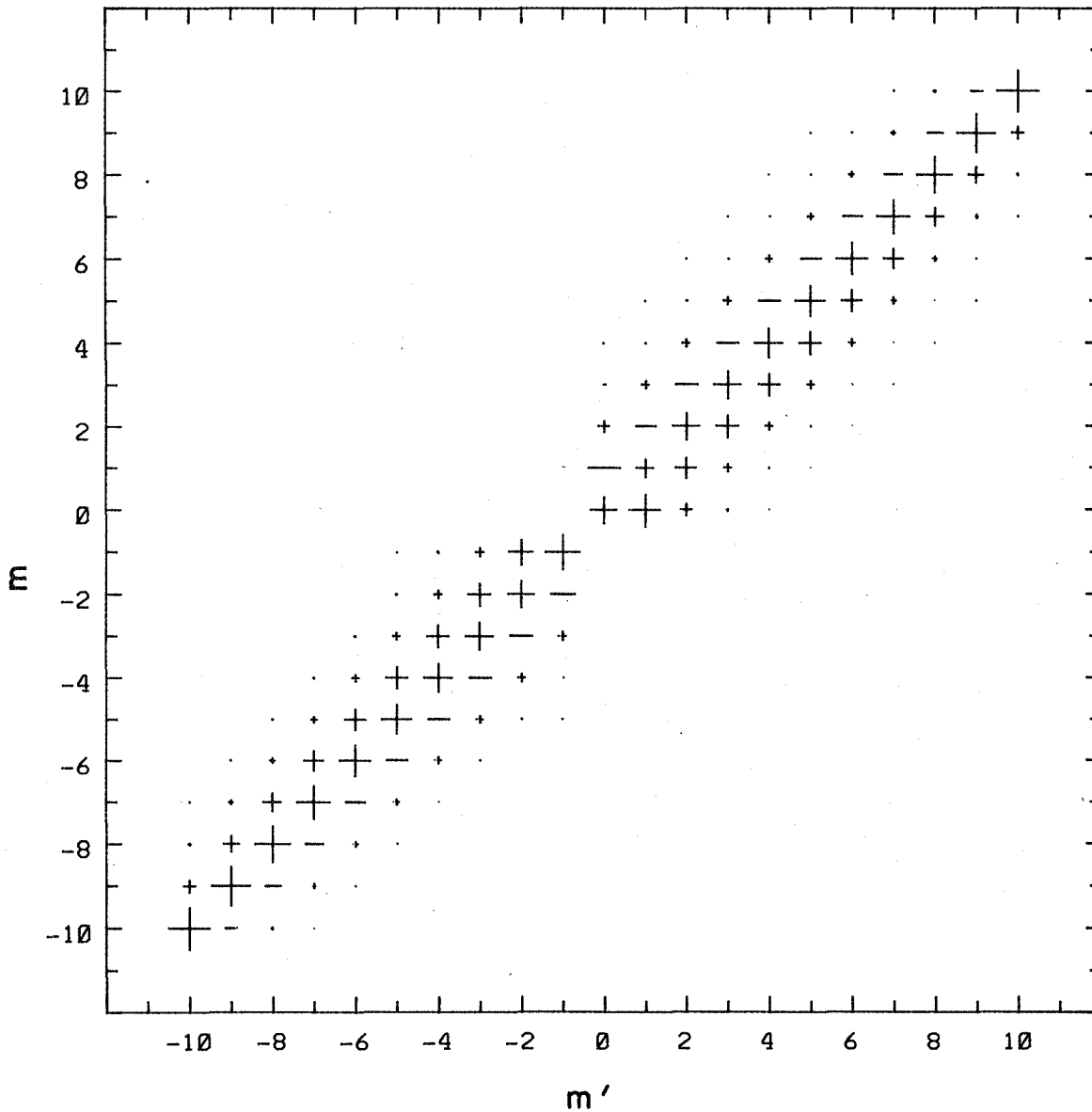


Figure 13. Transformation matrix  $\tilde{D}_{m',m}^{(\ell,t)}(\alpha, \beta, \gamma)$  for spatial responses, for  $\ell = 10$ ,  $\alpha = \gamma = 0$  and  $\beta = -7.25^\circ$  (corresponding to the maximum inclination of the rotation axis towards the observer). The format is the same as that used for the spatial responses (see the caption to Fig. 2).

rotation axis relative to the plane of the sky, defined to be positive when the North Pole is closest to the observer. Fig. 13 illustrates the transformation matrix  $\tilde{D}_{m',m}^{(\ell,t)}(0, -\theta_r, 0)$  for  $\ell = 10$  when  $\theta_r = 7.25^\circ$ , its maximum value. As might have been expected the matrix is diagonally dominated:

however there are significant matrix elements at  $m' - m = \pm 2$ . Thus even this fairly small inclination causes a substantial mixing in the observations of the modes.

While this gives rise to additional complication in the observed spectrum, the confusion is only in  $m$ ; thus the problems caused by frequency degeneracy in  $\ell$  are not seriously worsened. Furthermore one may correct for the mixing in the calculation of optimized combinations of the responses by using the rotated responses  $\tilde{S}'_{\ell m}(\mu)$  in the optimization procedure.

## Appendix B. The computation of pixel responses

To compute the pixel responses according to equations (3.2) and (3.3) we have to evaluate integrals over the pixels. This is done by quadrature, using Simpson's scheme in  $\theta$  and  $\phi$ . The mesh in  $\theta$  is the same for all pixels in a row, and the mesh interval is restricted to be less than an upper limit  $\Delta\theta_{max}$ . When the extent  $\Delta\theta_i$  in  $\theta$  of the pixel is below  $\Delta\theta_{max}$  the integral is replaced by the value of the integrand at the midpoint in  $\theta$  of the pixel, multiplied by  $\Delta\theta_i$ . Otherwise the mesh is uniform in  $\theta$ , with points on the edges of the pixels and an odd number of points on each pixel. The mesh in  $\phi$ , which is in general different for each mesh point in  $\theta$ , is defined similarly, the mesh interval being restricted to be less than  $\Delta\phi_{max}$ .

In the present calculation the pixels were set symmetrically with respect to the centre of the disk; hence the responses were only calculated for pixels in one quadrant and extended to the other quadrants using the symmetries of the spherical harmonics. We used  $\Delta\theta_{max} = \Delta\phi_{max} = 0.02$ . Comparisons with responses for selected  $\ell$  and  $m$ , calculated with  $\Delta\theta_{max} = \Delta\phi_{max} = 0.015$ , indicate that for  $\ell = 60$  the maximum errors in the pixel responses, in units of the rms of the responses over all the pixels, are less than 10 per cent, and the rms error is less than about 2 per cent. Thus the errors in the transformed responses are probably also at most a few per cent. The errors decrease rapidly with decreasing  $\ell$ .

The calculation of the Legendre functions requires a little care. We have used the recursion relation

$$(\ell - m + 1) P_{\ell + 1}^m(x) = (2\ell + 1)xP_{\ell}^m(x) - (\ell + m)P_{\ell - 1}^m(x) \quad (\text{B.1})$$

together with

$$P_m^m(x) = (-1)^m \frac{(2m)! 2^{-m}}{m!} (1-x^2)^{m/2} \quad (\text{B.2})$$

(e.g. Abramowitz & Stegun 1965). Thus for each  $m$ ,  $P_m^m$  is first evaluated from equation (B.2), and  $P_\ell^m$  is then calculated from equation (B.1) for increasing  $\ell$ , starting from  $\ell = m$ . This recursion appears to be stable: comparison of results obtained on computers with different word lengths indicates no serious loss of precision for  $\ell \leq 100$ . To avoid overflow it is advisable to work in terms of the scaled Legendre functions  $q_\ell^m P_\ell^m$  (cf. Section 2). It might also be pointed out that the organization of the response calculation fits in naturally with the recursion at fixed  $m$ : For a given value of  $m$  the integrals in equation (3.2) over  $\phi$ , at each of the mesh points in  $\theta$ , are independent of  $\ell$ : they are therefore evaluated only once, and only the integrals in  $\theta$  must be calculated at each value of  $\ell$  during the recursion.

It is quite likely that a recursive evaluation of the pixel responses, and hence of the transformed responses, can be found, similar to the one discussed in Appendix A of Christensen-Dalsgaard & Gough (1982). This deserves to be looked into: certainly the evaluation by quadrature would become prohibitively expensive in terms of computing resources if a substantially larger number of pixels were considered, or responses for higher values of  $\ell$  and  $m$  had to be computed.



## THE EFFECTS OF IMAGE MOTION ON THE $\ell$ - $\nu$ DIAGRAM

Frank Hill

National Solar Observatory\*, Sunspot, N.M. 88349 and  
Joint Institute for Laboratory Astrophysics,  
University of Colorado, Boulder, Colorado 80309

**ABSTRACT:** A simple two-dimensional  $(x,t)$  model of the solar oscillatory velocity field is subjected to a form of differential image motion. This image motion is meant to approximately model the effect of the Earth's atmosphere on observations of high degree solar oscillations. The distorted velocity field is analyzed to provide the apparent frequencies of the modes. Comparison of the results with the frequencies obtained from the undistorted case shows that the image motion can produce a discrepancy of as much as 12  $\mu\text{Hz}$ .

### I. Introduction

The goal of helioseismology is to measure the frequencies of the normal modes of the Sun as precisely as possible, and then to infer the structure of the solar interior from these observations. It is thus important to assess sources of noise in the observations of solar oscillations. In this paper, I report the results of a preliminary study of the effect of image motion produced by seeing in the Earth's atmosphere on the  $\ell$ - $\nu$  diagram of high degree solar oscillations.

Differential image motion redistributes power in the  $\ell$ - $\nu$  diagram by causing abrupt discontinuities in the original solar velocity images. These sharp features introduce frequencies other than the actual solar ones in the Fourier representation of the data. Image motion is thus a source of noise in the  $\ell$ - $\nu$  diagram, which may change the shape of the ridges, and could change the apparent centroid of the ridge as defined by Hill et al. (1984a, 1984b). This affects the results of the horizontal velocity inversion.

In this paper, I used a simple two-dimensional  $(x,t)$  model of the solar velocity field and subjected it to a form of image motion approximating the effects of the Earth's atmosphere. These artificially

---

\*Operated by the Association of Universities for Research in Astronomy, Inc. under contract AST 78-17292 with the National Science Foundation

distorted velocity fields were Fourier transformed to form  $\lambda$ - $\nu$  diagrams and then subjected to the same centroiding method described in Hill et al. (1984b). The deviation caused by the image motion was then measured.

## II. Method

The synthetic data was generated to simulate the observations we have made previously in our studies of subsurface horizontal velocities (Hill, Toomre and November 1982, 1983; Hill, Gough and Toomre 1984a, 1984b). The actual data was originally obtained with a total of two spatial dimensions ( $x, y$ ) and one temporal dimension ( $t$ ), and then averaged over one spatial dimension to form an ( $x, t$ ) array. The synthetic data was generated to simulate only the final state of the data; thus it was generated in two dimensions ( $x, t$ ). However, in order to model the transfer of velocities from pixel to pixel, the synthetic data was first calculated on a spatial grid of  $1/4''$ , a factor of 8 smaller than the  $2''$  resolution of the actual data. The grid was loaded with a total of 22 modes with  $n = 1, 2$  and  $\lambda = 500$  to  $550$  with a spacing  $\Delta\lambda$  of 5. Frequencies were generated from the formula relating  $n, \lambda$  and  $\omega$  found in Stein (1982). The modes were given constant and equal amplitudes and uniformly distributed random phases.

The image motion was simulated by starting from a mean two-dimensional ( $k, \nu$ ) Kolmogorov power spectrum, that is,

$$p(k) = \frac{k^{-5/3}}{k_0},$$

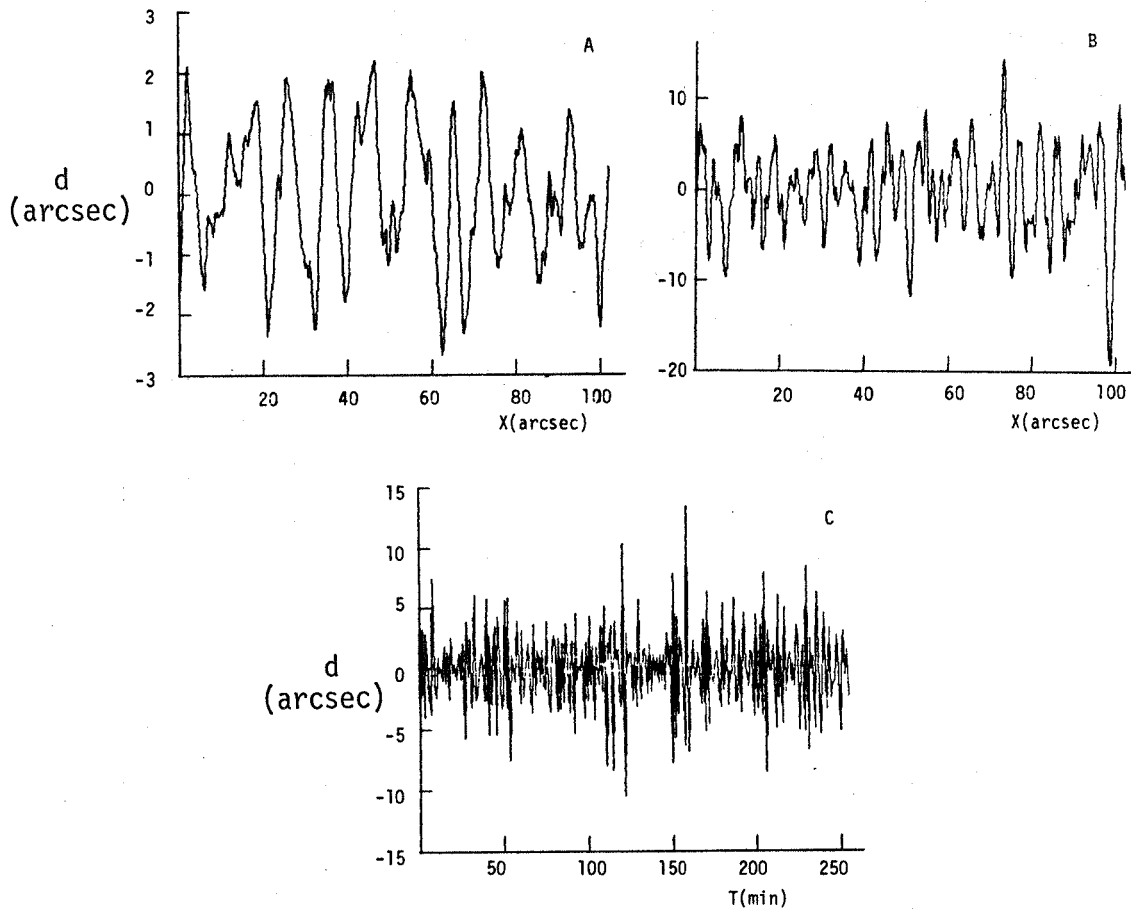
and a Gaussian spectrum

$$p(k) = e^{-(\ln k - \ln k_0)^2} \quad \text{for } 0 < k < k_0.$$

The parameter  $k_0$ , or cutoff wavenumber, was varied to provide a "coherence length" that is meant to provide the approximate size of a section of the image that moves as a unit. In all cases, the corresponding cutoff frequency  $\nu_0$  was set to 0.1 Hz, far above the temporal Nyquist frequency, thus providing a rapidly fluctuating temporal signature. The power spectral amplitude was given a random Gaussian distribution about the mean spectrum defined above. The phases were given a uniform random distribution between 0 and  $2\pi$ . The resulting two-dimensional power spectrum was then inversely Fourier transformed to produce an array of seeing displacements  $d(x, t)$ . The displacement array was scaled to give an RMS value defined to be the amplitude of the average seeing.

Some typical displacement patterns are shown in Figure 1. These plots are one-dimensional slices of the two-dimensional displacement array. Panels A and B are slices as a function of  $x$  at a given  $t$ ; panel C is a slice as a function of  $t$  at a given  $x$ . Panel A is for the case when the coherence length is set to  $10''$  and the amplitude is set to  $2''$ , and shows the approximately  $10''$  wavelength in the spatial pattern of

displacements. Panel B shows a case where the coherence length is 5" and the amplitude is 10". It shows a spatial wavelength of about 5". Panel C illustrates the aperiodic temporal behavior of the displacement resulting from the high cutoff frequency.



**Figure 1:** One-dimensional slices of typical seeing displacement arrays. Panel A: Displacement as a function of  $x$  at a given  $t$  for case 2. The amplitude is 2", the coherence length is 10". The 10" wavelength is readily apparent. Panel B: Displacement as a function of  $x$  at a given  $t$  for case 9. The amplitude is 10", the coherence length is 5". Again the coherence length is evident. Panel C: Displacement as a function of  $t$  at a given  $x$  for case 4. The amplitude is 10", the coherence length is 10". The aperiodic temporal character of the displacement is due to the high cutoff frequency.

The displacements were then applied to the data by using them to provide a map specifying where the velocity in a given pixel would appear in the "observed" data. The velocity in a pixel  $(x,t)$  was moved to a pixel  $(x',t)$ , where  $x'=x+d(x,t)$ , and then added to the existing pixel contents. Eight adjacent pixels were then averaged together to simulate observations covering  $512 \times 512$  points with 60 s time resolution and 2" spatial resolution. This final velocity field is then transformed and the resulting artificial ridge pieces subjected to our centroiding analysis. A total of nine cases were investigated in this study using various combinations of coherence length (L) and displacements (D). These nine cases are summarized in Table 1.

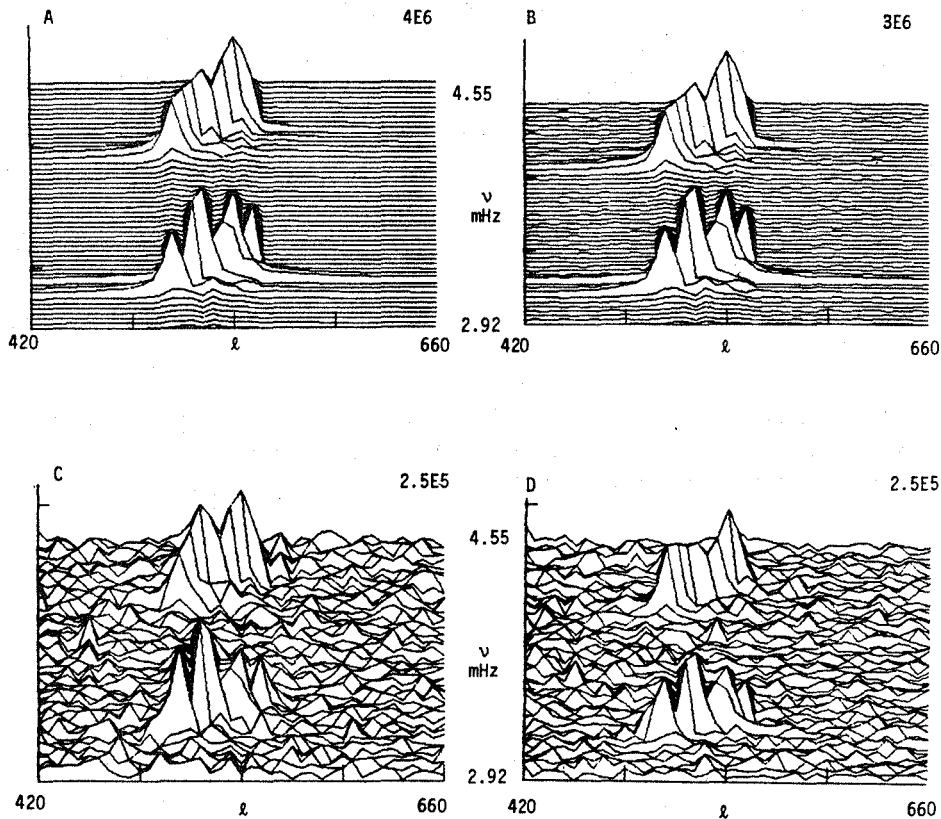
**Table 1**  
Image Motion Parameters

Case	Amplitude (arcsec)	Coherence Length (arcsec)
1	1	10
2	2	10
3	5	10
4	10	10
5	1	100
6	5	100
7	10	100
8	5	5
9	10	5

### III. Results

Of the nine cases considered, only cases 4, 7, and 9 produced any change to the ridge shapes. These 3 cases are the ones in which  $D=10''$ , close to the range of wavelengths of the modes of  $11.42''$  to  $12.57''$ . Figure 2 shows the relevant portion of the artificial  $l-v$  diagrams for cases 4 and 9, as well as the diagram of the undistorted data and the diagram of case 2, a relatively mildly distorted trial. The scaling of the individual plots has been adjusted to give about the same apparent ridge height in all four panels; the number in the upper right corner of each panel is a measure of the actual height. In case 7, which had a large coherence length of  $100''$ , the ridges were completely obliterated. At a displacement amplitude of  $10''$ , and coherence lengths of  $5''-10''$ , the shape of the ridges is substantially changed, although all of the main peaks seen in the undistorted case are also visible in every case of distortion. The background noise level is substantially raised, and the amplitudes of the ridges have dropped by a factor of about 16. For the "good seeing" case ( $D=2''$ ), the amplitude also drops, but by only 25%.

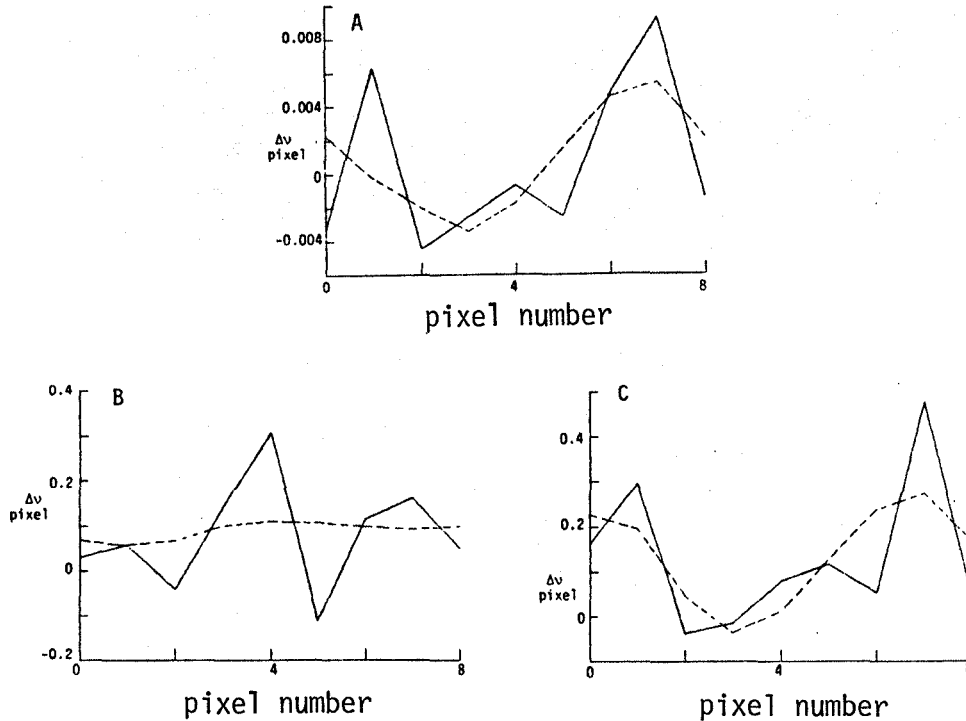




**Figure 2:** The  $l$ - $v$  diagram for some of the distorted data. Panel A: The undistorted diagram. The amplitude in arbitrary units is in the upper right corner. Panel B: The diagram for case 2, a mild distortion with an amplitude of 2" and a coherence length of 10". Very little change is evident and the amplitude has dropped by 25%. Panel C: The diagram for case 4, a more severe case with an amplitude of 10" and a coherence length of 10". The shape of the ridges is greatly altered, the relative background noise level is substantially higher and the ridge heights have dropped by a factor of 16. Panel D: The diagram for case 7, also a severe case with 10" amplitude and 5" coherence length. The results are similar to panel C.

This increased background noise level and decreased ridge amplitude produce an error in the determination of the centroid and hence in the measured frequencies of the modes. The artificial ridge segments were analyzed by the same centroiding and fitting procedure used previously (Hill, Toomre and November 1982, 1983) and discussed more fully in Hill, Gough and Toomre (1984b). For this artificial case, the ridges were cut in the vertical (constant  $l$ ) direction. A 3 knot fit was run through the resulting centroids. The resulting ridge positions were then com-

pared to the ridge positions obtained from the undistorted case. Figure 3 illustrates the magnitude of the difference between the undistorted and the distorted cases as a function of position along the "p1" ridge. Both the difference and the position are measured in pixels. The bold curve is the difference in the fit, the dashed curve is the difference in the raw centroids. Panel A is for the  $D=2''$ ,  $L=10''$  case, and it can be seen that the difference is only a few millipixels. For this data, 1 pixel =  $30.7 \mu\text{Hz}$ , so the error here is about  $0.15 \mu\text{Hz}$ . Panel B shows the results for the  $D=10''$ ,  $L=10''$  case, and panel C is for the  $D=10''$ ,  $L=5''$  case. For both of these cases, the difference from the undistorted case is from 0.1 to 0.4 pixels, or about 3 to 12  $\mu\text{Hz}$ . The difference in the fitted centroids is typically smaller than the difference in the raw centroids. The same plots are shown in Figure 4 for the "p2" ridge, with similar results. The corresponding error in velocity for the modes shown here is from 25 to 100 m/s. These errors are substantial, and even though the inversion procedure greatly reduces the errors, they remain a significant source of noise.



**Figure 3:** The difference in the centroids of the "p1" ridges seen in Figure 2 as a function of position along the ridge. All differences are relative to the centroids of the ridges in Figure 2A, the undistorted case. The solid lines are the difference in the fitted centroids, the dashed lines are the differences in the raw centroids. All units are in pixels, with 1 pixel in  $\Delta v$  corresponding to  $30.7 \mu\text{Hz}$ . Panel A: Difference between case 2 and the undistorted diagram (Figures 2A and 2B). Panel B: Difference between case 4 and the undistorted diagram (Figures 2A and 2C). Panel C: Difference between case 9 and the undistorted diagram (Figures 2A and 2D).

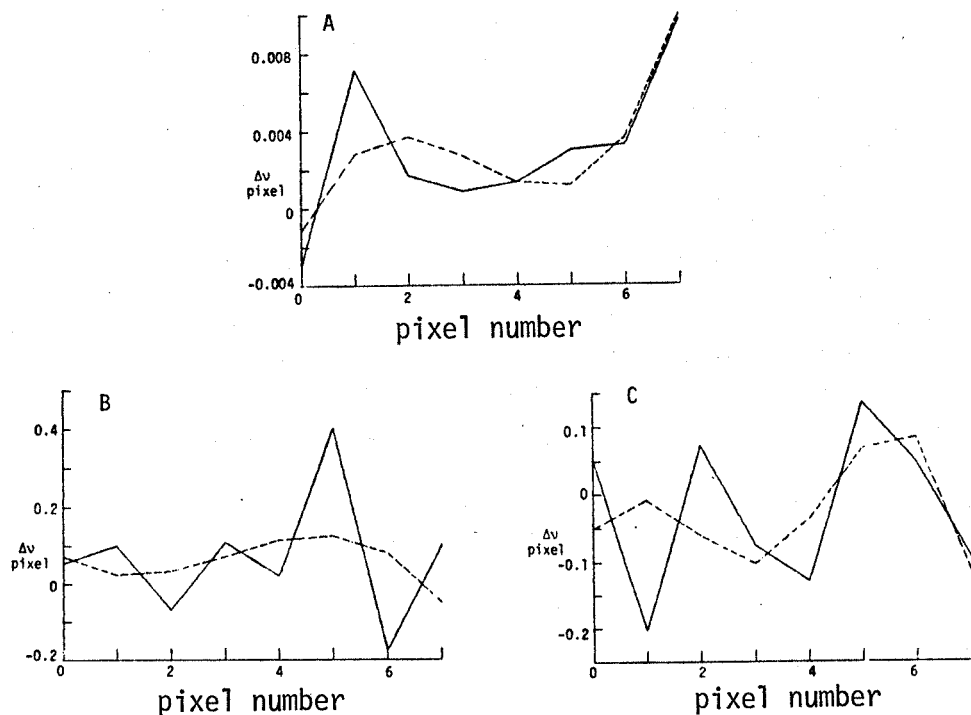


Figure 4: As Figure 3, but for the "p2" ridge.

#### IV. Discussion

It is not surprising that image motion severely affects the  $l$ - $v$  diagram only when the displacement amplitude is on the order of the wavelength of the modes. It is well known from many physical situations that waves are generally not affected by background disturbances with characteristic spatial scales that are either much smaller or much larger than the length of the waves, but are disrupted when the background pattern has a spatial scale on the order of a wavelength. The same qualitative effect apparently holds in this study. The 10" seeing that causes the most disruption would typically be considered quite poor at a ground-based observatory; however, better seeing conditions would probably affect modes with higher values of  $l$  and thus shorter wavelengths in a manner similar to that seen in this study. One might expect the  $l$ - $v$  diagram to be adversely affected only in the regions corresponding to the amplitude of image motion prevalent at the time of observation.

The model presented here is crude and has a number of limitations. First, the model has only one spatial dimension, whereas actual data generally has two spatial dimensions. The addition of the second spatial dimension may affect the mode cancellation properties of any filtering performed on the data, and may introduce different frequencies from other modes as well. Secondly, no background steady velocity

fields such as supergranulation have been added. Such a velocity field may be a substantial noise source when image motion is considered (Ulrich 1984). Thirdly, the "observations" were assumed to be instantaneous, rather than integrated over a finite length of time. A longer "exposure" time may help to average out some of the effects of image motion. Fourthly, the model considers velocities only, and does not take the nonlinearities of spectral line shape or detector response into account. Fifthly, only a small number of modes and consequently short pieces of ridges were considered. The addition of more modes covering a larger range in  $\lambda$  would presumably show that different segments along the ridges are affected as the amplitude of the image motion approaches the wavelengths of the modes. Finally, the seeing was assumed constant rather than variable throughout the observational span.

In addition, the original synthetic data was calculated with a mode spacing that is much wider than in reality. With a spacing of  $\Delta\lambda = 5$ , there was only about 1 mode in each resolution bin in the resulting  $\lambda$ - $\nu$  diagram. The actual spacing in  $\lambda$  and  $\nu$  of the modes is such that there are actually approximately 30 modes per bin. The beating between such a large number of modes is another factor that may redistribute power in the ridges and may thus be a source of error in inversions. This effect has not been considered here.

In conclusion, it appears that image motion caused by the Earth's atmosphere may be an important source of error in the inversion of solar oscillation data obtained from ground-based observations. However, further more detailed modelling must be done to ascertain whether the error so introduced is severe enough to warrant a satellite borne helioseismology instrument. In addition, image motion must be taken into account when developing procedures to combine the data from different stations in a ground-based network.

**ACKNOWLEDGEMENTS:** It is a pleasure to thank Douglas Gough for the suggestion of a displacement mapping to model the seeing, and to thank Juri Toomre, Roger Ulrich, Jack Harvey, Tim Brown and Bill Merryfield for useful discussions. Part of this work was done at the University of Colorado and was partially supported by the Air Force Geophysical Laboratory under contracts F19628-77-C-0104 and F19628-83-K-0008 and by NASA through grants NSG-7511 and NAGW-91 to the University of Colorado.

## References

- Hill, F., Gough, D.O., and Toomre, J.: 1984a, submitted to Proceedings of Conference on Oscillations as Probes of the Sun's Interior, Catania, Italy.
- Hill, F., Gough, D.O., and Toomre, J.: 1984b, these Proceedings.
- Hill, F., Toomre, J., and November, L.J.: 1982, in Proceedings of Conference on Pulsations in Classical and Cataclysmic Variable Stars, ed. by J.P. Cox and C.J. Hansen, J.I.L.A., University of Colorado, p.139.
- Hill, F., Toomre, J., and November, L.J.: 1983, Solar Phys., **82**, 411.
- Stein, R.F.: 1982, Astron. and Astrophys., **105**, 417.
- Ulrich, R.K.: 1984, these proceedings.

## THE EFFECTS OF SEEING ON NOISE

**Roger K. Ulrich**

Department of Astronomy, University of California, Los Angeles,  
CA 90024, U.S.A.

**Edward J. Rhodes Jr.**

Department of Astronomy and Earth and Space Science Institute,  
University of Southern California, Los Angeles, CA 90007 and  
Jet Propulsion Laboratory, California Institute of Technology,  
Pasadena, CA 91109, U.S.A.

**Alessandro Cacciani**

Dipartimento di Fisica dell-Universita "La Sapienza" Roma, Italy.

**Steven Tomczyk**

Department of Astronomy, University of California, Los Angeles,  
CA 90024, U.S.A.

**Abstract:** We discuss the effect of the supergranulation velocity field combined with seeing smearing of the solar image on the measurement of solar oscillations. Depending on the nature of the observational velocity determination scheme, the image motions can shift the background velocity pattern and produce a source of noise that reduces the quality of the observations. We give a rough estimate for the magnitude of this effect and present observational results which are consistent with this estimate.

### **1. Introduction**

The presence of turbulence in the earth's atmosphere causes the image of the sun to be distorted and shifted in a fashion which depends randomly on time. Part of the distortion moves the image as a whole and is removable by guiding during the observation or by registration after the observation. Another part of the distortion which occurs during an exposure causes the image to be smeared. Unfortunately, a large part of the seeing effect causes the image to be distorted internally making it impossible to register or guide on all parts of the image simultaneously. Since the background velocity associated with both the supergranulations and the oscillations themselves has large transverse gradients, the shifting of the image from one pixel to another can produce an error in the measured velocity. One application of the oscillation measurements will be to determine the frequencies of the larger scale modes. In this case the shifting of the velocity from one pixel to another may cancel out and leave the result better than a simple estimate would indicate. In practice the degree of such cancelation is difficult to predict because of such effects as pixel to pixel variations in sensitivity, minor non-linearities in the velocity determination scheme, and sub-pixel structure like gaps between pixels. These imperfections will prevent perfect cancelation even for the largest scale modes.

For the modes with structure near the scale of the background variations, there is no cancelation even for a perfect observing system.

Recent measurements by Forbes (1982) of seeing induced motions of the binary star  $\gamma$  Andromeda showed that if the motion of one component were followed then the motion of the second component only  $9.6''$  away would have been reduced by just a factor of 2. Consequently, we cannot rely on image registration to compensate for seeing distortion. The problem in the case of solar oscillation measurements comes from the need to cancel the large velocities of the solar supergranulation very precisely. Some observing systems such as the full disk resonance cell technique used by Grec, Fossat and Pomerantz (1980) and Claverie, et al (1979, 1981), and the optically averaged system used by Duvall and Harvey (1983) avoid the need to have accurate cancellation of the light shifted from one pixel to another because the detecting pixel is much larger than the supergranulation scale. However, such systems have limited ability to resolve the exact spatial structure of the modes of oscillation and consequently cannot fully determine the dynamical state of the interior of the solar convective envelope. In contrast, when the pixel size is smaller than the supergranules, image motion will cause the average velocity of the surface observed by the pixel to change by an amount which is equal to the product of the image displacement and the velocity gradient within the supergranule. Since a fully resolved observing system will probably use pixels a few arc-seconds on a side and the supergranulation has a scale of roughly  $40''$ , we can assume that the pixel is smaller than the supergranule. Thus we treat the change in the pixel velocity as just the result of the shift in the position of the pixel relative to the supergranule. Some fraction of this velocity change will not be cancelled and will show up as a velocity error.

In order to test these ideas, we have analyzed observations obtained at Mt. Wilson with the Magneto-Optical Filter on July 27, 1983. These observations are described in detail elsewhere in this volume. In this paper, we have broken the full day's observing sequence into morning and afternoon halves on the assumption that the quality of the seeing degrades during the day. The two halves were reduced separately and compared to each other and to the combined spectrum. This comparison is described in section III below.

## 2. Velocity Noise Due To Supergranulation

The velocity gradient associated with the supergranulation is given approximately by

$$\frac{dv}{dx} = \frac{2 v_0 \sin \theta}{L/2 \cos \theta} \quad (1)$$

where  $v_0$  is the peak horizontal velocity of the supergranulation,  $x$  is the position on the apparent solar image measured in arc-seconds, and  $L$  is the center to center distance between cells also in arc-seconds. For numerical work we use the results of Simon and Leighton (1964) and take  $L = 32''$  and  $v_0 = 420$  m/s. A more precise analysis which represents the supergranulation as a two dimensional planform would yield an RMS velocity within about a factor of

$2^{1/2}$  of the value given by equation (1). The angle  $\theta$  is the center to limb distance. Seeing causes random motions of the image position. Various theories are available to describe these motions with the simplest being the Kolmogorov described by the theory of turbulence. The work by Forbes shows a frequency dependence of the image motions which is consistent with the Kolmogorov theory. Other models of seeing such as that described by Brunner(1982) have slightly different frequency dependence. The effect of using a different exponent will be minor for our application. Adopting the simple Kolmogorov model we take the power spectrum of the displacement of the apparent position of a fixed part of the sun to be

$$P_s = \frac{a}{\nu^{5/3}} \quad (2)$$

where  $P_s$  is the power spectral density (measured in arc-seconds<sup>2</sup> / Hz) and  $a$  is an amplitude which is proportional to the commonly quoted seeing size  $s_0$ . In order to evaluate  $s_0$  we must integrate  $P_s$  over some range in frequency. The seeing size normally is taken to mean the size of the blurred image of a point source. All frequencies above some guiding frequency contribute to this blur. Thus we integrate  $P_s$  from  $\nu_0$ , the guiding frequency, to infinity. The resulting expression relating  $a$  to  $s_0$  is then

$$a = \frac{2}{3} s_0^2 \nu_0^{2/3} \quad (3)$$

Although expressions such as that given by Brandt(1969) provide estimates of  $a$  based on factors like the wind velocity and telescope size, we prefer to leave the normalization in the form of equation (3) so that the image blur appears explicitly in our subsequent equations. The total velocity noise generated by seeing  $\Delta v^2$  is the integral from the sampling frequency  $\nu_s$  to infinity of  $P_s$  times the square of the velocity gradient. After carrying out this integral we find

$$\left( \frac{\Delta v}{840 \text{ m/s}} \right)^2 = \left( \frac{s_0}{16''} \right)^2 \left( \frac{\nu_0}{\nu_s} \right)^{2/3} \tan \theta \quad (4)$$

For values of  $\nu_s$  below  $\nu_0$ , the effect of the seeing variations can be partially compensated by the guiding system. The coherence of the image tends to increase as  $\nu$  decreases so that the guiding at low frequencies is more effective in compensating for the seeing motion. Thus for low frequencies we assume that  $(\nu_0/\nu_s)^{2/3}$  is replaced by a compensation factor  $C(\nu_s)$  which describes the effectiveness of the guiding system and the fraction of the image motion which is internal distortion. The results of Forbes would suggest that  $C(\nu_s)$  should be of order 0.5 when the guider works well whereas the work by Fossat, Grec and Harvey (1981) indicates that the internal distortion for the low frequency variations is on the order of 0.01 times the image motion.

Numerically equation (4) becomes:

$$\Delta v^2 = (50 \text{ m/s})^2 C(\nu_s) (s_0)^2 \tan \theta \quad (5)$$

Both the solar velocity signal and the seeing induced noise are averaged over space and time when they are reduced to power on a  $\lambda - \nu$  plane. Before the reduction, the amplitude of the seeing induced noise given by equation (5) should be compared to  $v_{os}$ , the integrated amplitude of the five minute oscillations, which is 500 m/s. Because the supergranulation structure is spatially irregular and the high temporal frequency part of the seeing distortion is enough to introduce large velocity changes, the frame to frame positional correlation implied by the low temporal frequency part of the seeing motion should not cause a frame to frame correlation in the velocity variations. Consequently, the power spectrum of the seeing induced velocity error should be independent of temporal frequency and spatial wave-number. Without additional information about the distribution of the seeing induced noise, we assume that it is distributed uniformly in frequency and wavenumber. The amplitude of the oscillatory power relative to the seeing induced noise then depends on the distribution of the solar power over the  $\lambda - \nu$  plane. In particular if the solar power is concentrated into a fraction  $F$  of all the points on the  $\lambda - \nu$  plane then the relative noise due to the seeing will be reduced by this factor. One part of  $F$  comes from the fact that the five minute oscillations are concentrated into a band about 1500  $\mu\text{Hz}$  wide. This introduces a factor of  $1500/\nu_s$  into  $F$ . Another factor comes from the fact that the power is concentrated toward low values of  $\lambda$ . In our spectra, the power is confined to the lower quarter of the  $\lambda$  range. The further concentration of power to the eigenfrequency ridges within the high power regions of the  $\lambda - \nu$  plane depends on several details such as the lifetime of the modes, the resolution in  $\lambda$ , and the derivative of  $\nu$  with respect to  $\lambda$ . In our spectra which are for a single day's observation and which are confined to the central 40% of the solar disk, the concentration to the ridges adds about another factor of one third to  $F$ . Combining these factors we find that the ratio of the seeing noise to the five minute signal is

$$\left(\frac{\Delta v}{v_{os}}\right)^2 = 0.01 C(\nu_s) F s_0^2 \tan \theta \quad (6)$$

For our observations the value of  $C$  is probably about 0.1 and the value of  $F$  is about 0.01 so that the seeing induced noise level is given roughly by:

$$\Delta v^2 = 10^{-5} v_{os}^2 s_0^2 \tan \theta \quad (7)$$

Although the level of noise due to this source is small relative to the stronger peaks within the five minute band, for weaker oscillations at other frequencies, the seeing noise is a significant factor as we will show below.



Due to the dependence of noise on the center to limb angle, the noise power must be averaged over the solar disk in a weighting integral of the type described by Christensen-Dalsgaard and Gough (1982). However, in contrast to the case considered by Christensen-Dalsgaard and Gough, the limb darkening function is not included in the integral for the spatially resolved average because the determination of the velocity is done on a point by point basis.

The averaging integration introduces a factor of  $(-\ln(1 - b^2) / b^2 - 1)$  where  $b$  is the radius of the observed disk measured in units of  $R_{\odot}$ . This factor is only 0.47 if the observed portion of the solar disk extends 75% of the distance from the center to limb. It, however, the observed portion is 90% or 95% of the distance to the limb then the factor increases to 1.0 and 1.5 respectively. In our observations,  $b$  is 0.35 and the averaging factor is 0.085.

### 3. Observations

As described elsewhere in this volume, our data set consists of a 9.1 hour sequence of velocity measurements made with the magneto-optical filter described by Cacciani and Rhodes. The doppler-gram frames were obtained every 32 s during this period and each frame consists of a 128 x 128 grid of pixels. The image scale was approximately  $3.9 \text{ m}/\text{pixel}$ . Because of the experimental state of the setup on that day, the precise positions of all optical components were not recorded and we have been unable to reconstruct them subsequently. Consequently, the image scale is accurate to only 10%. The full day's data set consisted of 1024 frames. We divided the set into two subsets of 512 frames and reduced both subsets and the full set in the same way. We were careful to maintain a consistent normalization for all three power spectra so that the effects of the anticipated seeing degradation in the afternoon could be studied by comparing the strength of the oscillations as well as the strength of the background noise. Because of the relatively limited length of our spatial baseline, our resolution in  $\lambda$  produces a pixel which is wider in the  $\lambda$  dimension than in the  $\nu$  dimension when the data are displayed in a way which adequately illustrates the eigenmode structure. Figure 1 shows the three power spectra adjacent to each other. The combined spectrum is also given in a larger format and compared to theory in the paper by Rhodes et. al. elsewhere in this volume. The density of dots in Figure 1 is proportional to the rms velocity amplitude rather than the power.

One feature is immediately obvious from figure 1 - the eigenfunction ridges extend to higher  $\lambda$  values for the morning half relative to the afternoon half. The degradation in the apparent power begins at about  $\lambda=450$ . Since this value of  $\lambda$  corresponds to a wavelength of about  $10^{\hat{u}}$ , it is a bit surprising that the ridges should be degraded at so small a value of  $\lambda$ . Although we do not have a quantitative way of measuring the seeing, we believe that it is typically  $3^{\hat{u}}$ . Distortions of the wave pattern at a scale of about one third the wavelength may be sufficient to broaden the ridges until they effectively disappear. Our data indicate that the observability of the ridges can be easily degraded by seeing at moderate values of  $\lambda$ . Figure 1 also shows that the power outside the five minute band is increased in the afternoon relative to the morning. This effect is not as obvious as the ridge degradation and requires a more quantitative analysis to bring it out in a clear fashion. In order to smooth out the random variations in the power associated with the chaotic nature of the oscillatory

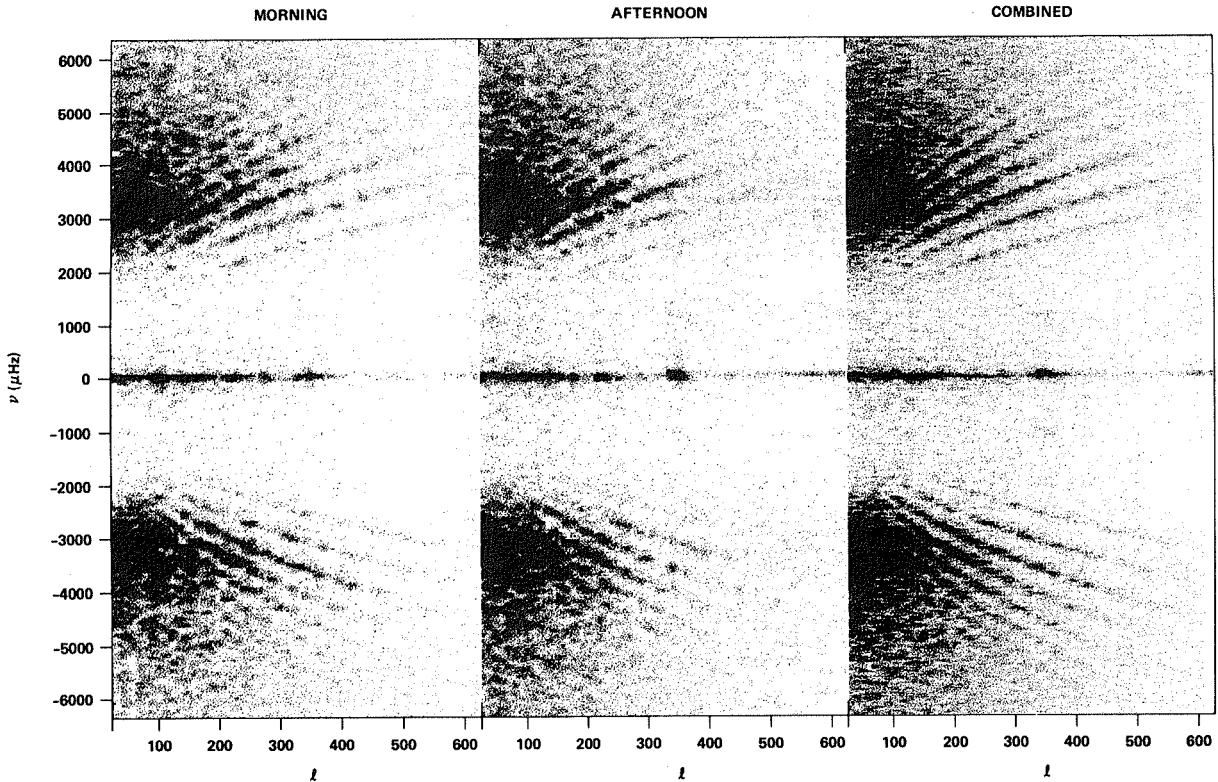


Figure 1. A comparison of the power spectral density in the morning to the power spectrum in the afternoon and the power spectrum for the full day. The density of points is proportional to the amplitude of the motion and is adjusted to be the same in all three panels.

motion, we averaged the power over moderately large rectangular sections of figure 1. Each rectangular area extended over 485  $\mu\text{Hz}$  in frequency and had end points in  $t$  as indicated in figures 2 and 3. The ratio of the power in the morning half to the power in the afternoon half was then calculated for each rectangle and plotted in figures 2 and 3 as a function of  $\nu$ . The arrows at the bottom of the figures indicate the positions of the peak power in the five minute band. The lower power in the morning for the very high and low frequencies where we expect the signal of solar origin to have a lower amplitude is strongly suggestive of the presence of a white noise source of the type we believe could be caused by seeing effects. In units which are dependent on the reduction process, the peak solar oscillation amplitude is 19. The amplitude of the oscillations where there may be an extraneous noise source is about 0.3 in the same units. For  $s_0 = 3$ , equation (7) gives  $\Delta\nu = 0.06$  in the same units. This result is a bit smaller than suggested by figures 2 and 3 but considering the many uncertainties in the analysis, seeing is a plausible cause for the noise.

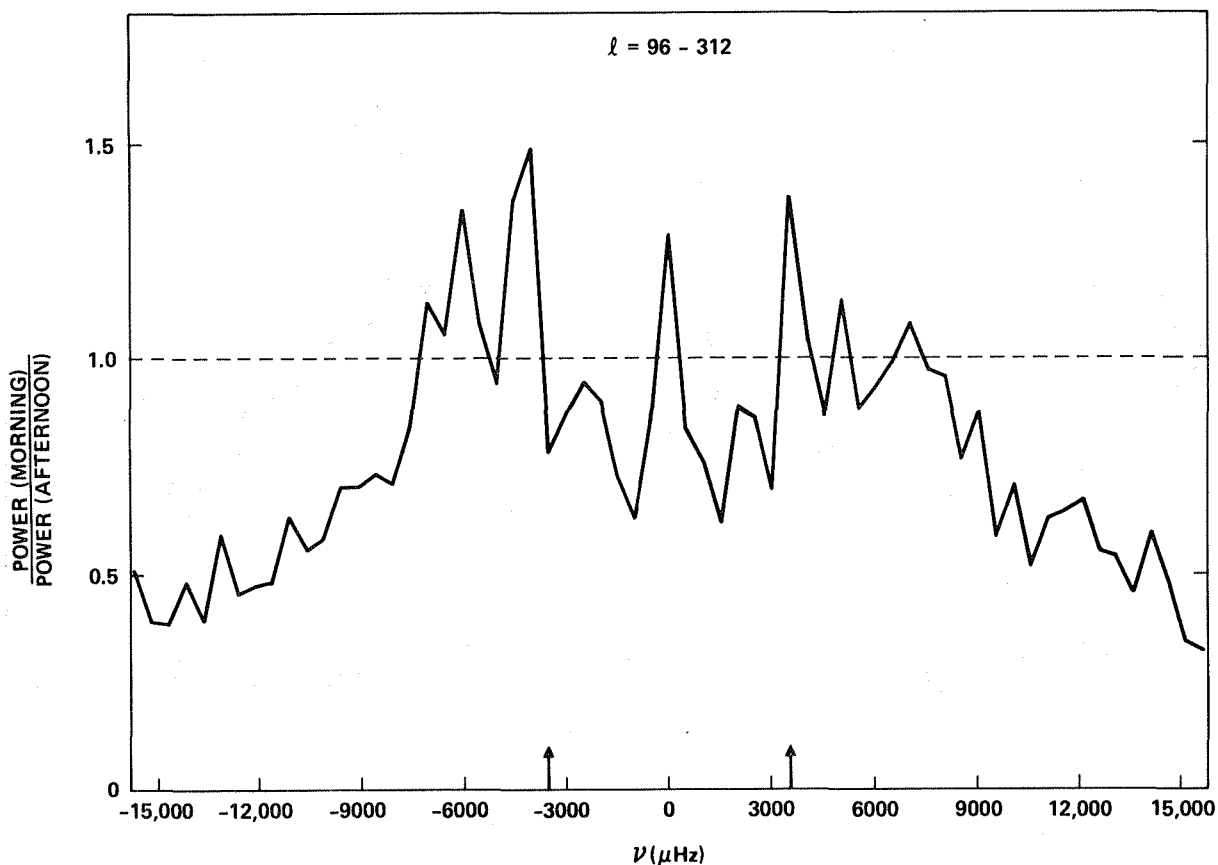


Figure 2. The ratio of the power in the morning to the power in the afternoon. Both morning and afternoon power was averaged separately over rectangular areas spanning the indicated range in  $l$  and a range of  $600 \mu\text{Hz}$ .

#### 4. Conclusions

The effect of seeing on observations of the solar oscillations was shown to be a plausible cause for the degradation of the power spectrum at high  $l$  and for the introduction of a background, broadband noise. For the observation of the stronger oscillations, the background noise will not have a major impact. However, the loss of the low amplitude oscillations into this background especially at long periods could seriously limit the seismic analysis of the solar interior. The longer period oscillations are known to have sharper frequencies which potentially can provide the most accurate constraints on the interior structure and dynamics. The low amplitude of these modes makes them especially vulnerable to the seeing generated noise.

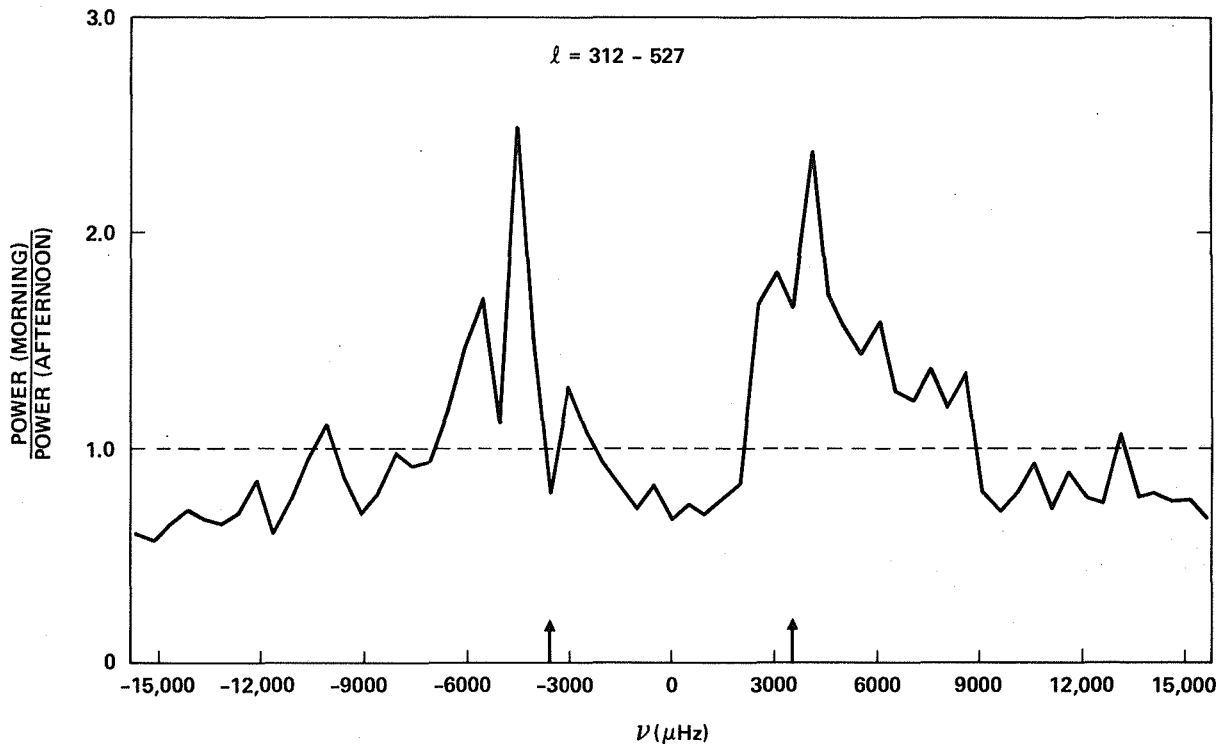


Figure 3. The same ratio as in Figure 2 but for the indicated higher range of  $l$ .

### References

- Brandt, P.N. 1969, *Solar Physics*, 7, 187.
- Brunner, F.K. 1982, *Sun and Planetary System*, W. Fricke and G. Teleki eds., D. Reidel, p. 505.
- Christensen-Dalsgaard, J. and Gough, D. O. 1982, *Mon. Not. R. Ast. Soc.*, 198, 141.
- Claverie, A., Isaak, G.R., McLeod, C.P., van der Raay, H.B., and Roca Cortes, T. 1979, *Nature*, 282, 591.
- , 1981, *Nature*, 293, 443.
- Duvall, T.L. Jr., and Harvey, J.W. 1983, *Nature*, 302, 24.
- Fossat, E., Grec, G., and Harvey, J.W. 1981, *Astron. Astrophys.* 94, 95.
- Forbes, F. 1982, *SPIE*, 332, 185.
- Grec, G., Fossat, E. and Pomerantz, M. 1980, *Nature*, 288, 541.
- Simon, G. W. and Leighton, R. B. 1964, *Astroph. J.*, 140, 1120.

## THE EFFECTS OF A NEARLY 100% DUTY CYCLE ON OBSERVATIONS OF SOLAR OSCILLATIONS

**Frank Hill**

National Solar Observatory\*

Sunspot, N.M., and

Joint Institute for Laboratory Astrophysics

University of Colorado.

**ABSTRACT:** Power spectra of window functions with duty cycles between 80% and 99% and with randomly spaced gaps are computed and their effect on observations of solar oscillations are discussed. It is found that for all the cases considered, observations of solar oscillations would not be severely impacted as long as the gap structure is random rather than periodic.

### **I. Introduction**

Currently, the major obstacle to a full helioseismological inversion is the difficulty of resolving and identifying modes in the dense spectrum of solar oscillations. The fact that almost all ground-based observations obtained at a single site are subject to the day/night cycle greatly complicates the task. This cycle produces strong side-lobes centered on each solar frequency, which can overlap and mask other solar frequencies (e.g. Brown 1979). This has proven to be a problem in identifying the splitting of frequencies due to differential rotation (Claverie et al. 1981) and in the identification of long period oscillations (Delache and Scherrer 1983, Bos and Hill 1983). Observations obtained at the South Pole during the Austral summer are free from this cycle but have a maximum duration of only about five days (Grec, Fossat and Pomerantz 1983; Stebbins and Wilson 1983; Harvey, Pomerantz and Duvall 1982). These observations provide power spectra with a frequency resolution of about 2  $\mu\text{Hz}$ , substantially higher than the few tenths of a  $\mu\text{Hz}$  desirable for helioseismology. A network of ground-based stations, or space-based observations from a satellite with a continual view of the Sun, is required to obtain both high frequency resolution and a duty cycle as close to 100% as possible. However, even in these cases a 100% duty cycle will be extremely difficult to achieve. Here, the duty cycle is defined as the fraction of time that the Sun is visible. The purpose of this paper is to investigate the effects of a nearly 100% duty cycle with randomly spaced gaps on the observations of solar oscillations.

---

\*Operated by the Association of Universities for Research in Astronomy, Inc., under contract AST 78-17292 with the National Science Foundation

Such duty cycles may be what we can expect to obtain from network or satellite observations.

## II. Method

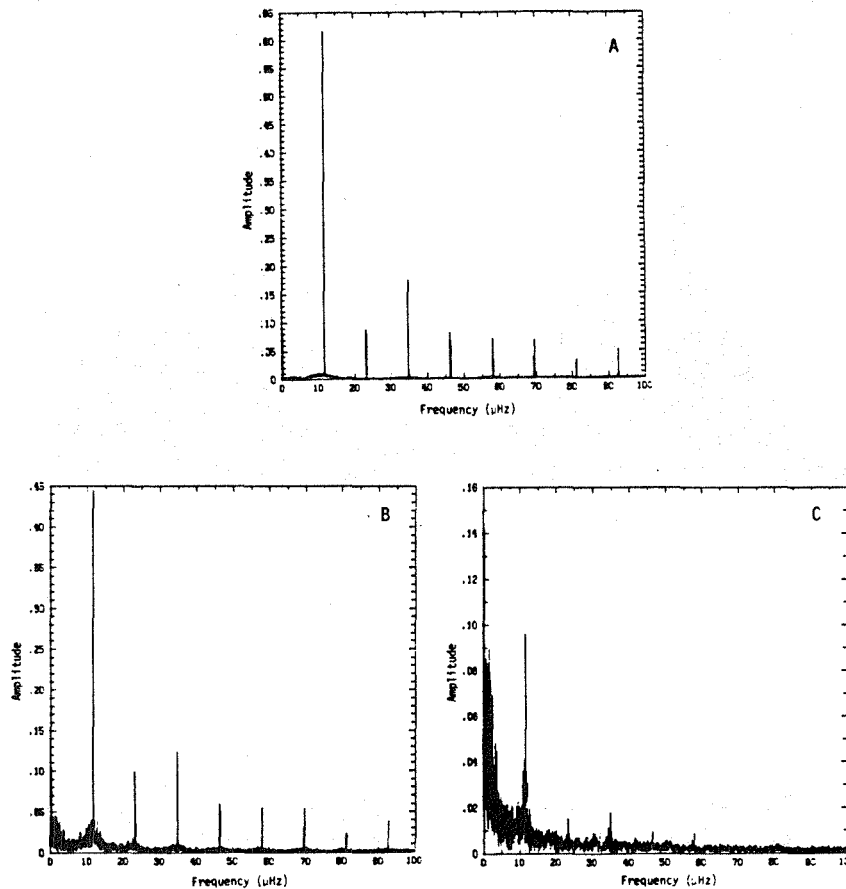
Actual observations result in a time series of velocities that is interrupted during various intervals. This can be represented in the real domain by the product of the actual solar signal and a window function that is 1 when the data is being obtained, and 0 otherwise. In the frequency domain, the observed power spectrum is the convolution of the actual solar spectrum and the spectrum of the window function, neglecting any other instrumental, atmospheric or solar sources of noise. Thus, every real peak in the solar spectrum will be surrounded by an image of the window spectrum. The method used in this study was to simply generate a number of windows with duty cycles of 80% to 99% and then to compute the power spectra of the windows. The gaps in the window function were generated by a number of methods. To simulate the window at a single ground-based site, an ephemeris equation was used to compute the approximate rising and setting times of the Sun at Sacramento Peak Observatory. In addition, a crude model of the weather throughout the year was produced from the author's memory of the seasonal weather patterns. These patterns were basically alternating periods of clear and cloudy weather lasting from 1 to 7 days during the months of November to April, and of clear mornings and cloudy afternoons from May to October. To generate window functions with randomly spaced gaps, the duty cycle  $D$  and average length of the gaps,  $L_g$  were first specified. This fixes the length of time between gaps,  $L_b$ , by

$$L_b = \frac{DL_g}{1-D}$$

Each individual gap had a uniformly distributed random length of between  $0.5 L_g$  and  $1.5 L_g$  and was separated by a uniformly distributed random length  $L_b$  between  $0.5 L_b$  and  $1.5 L_b$ . The window was generated with a 1 min time sampling over a total length of 1 year (525,600 samples); thus the power spectrum had a frequency resolution of 31.7 nanoHz.

## III. Results

Figure 1 shows the window function resulting from the day/night cycle and the weather at a single ground-based site (SPO). In this and all subsequent figures, the ordinate is the amplitude relative to the central ( $\nu = 0$ ) peak and represents the power of the surrounding window sidelobes relative to any central solar frequency. Figure 1A is the day/night cycle alone, showing the strong periodic structure at 11.57  $\mu\text{Hz}$  (1/day) and its harmonics. It is these sidelobes, particularly the first one which has an amplitude of over 60% of the central peak, that cause the severe confusion complicating mode identification. It should be noted that any periodic gap structure will introduce strong sidelobes at the gap frequency and its harmonics. Figure 1B shows the day/night cycle with weather included. The weather adds additional noise at the

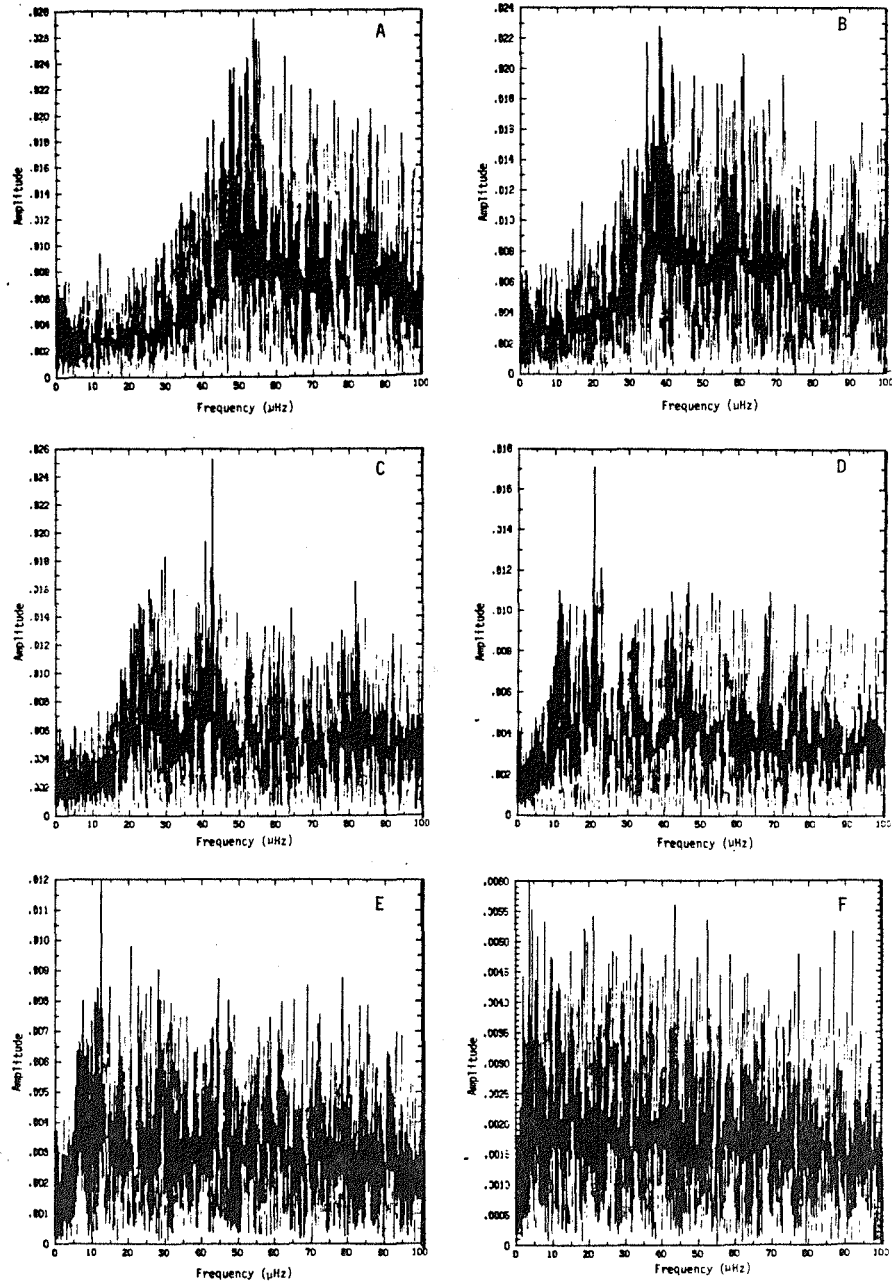


**Figure 1:** Power spectra of the window function that would result from observations at Sacramento Peak Observatory. A: Day/night cycle throughout year. B: Day/night cycle with approximate weather patterns throughout year superimposed. C: Weather patterns alone. Amplitude is relative to central peak in this and all other figures.

5% level and further reduces the duty cycle. Figure 1C is the transform of the weather alone. It also shows the day/night cycle due to the modelling procedure.

Figure 2 shows the window power spectrum for six different duty cycles of 80, 85, 90, 95, 97 and 99%, all with average gap lengths of 1 hr. Such a gap length might be typical of the observations obtained by a seismology instrument on board a satellite carrying a number of other instruments with incompatible pointing requirements. The spectra have a dense "grass" structure that peaks at a frequency determined by the average total length of time between gap beginnings ( $=L_g/(1-D)$ ). The amplitude of the peak of noise is quite low, being about 1.5% for the 80% duty cycle and dropping to about 0.2% for the 99% duty cycle. Thus, it appears that all of these duty cycles would be suitable for helio-

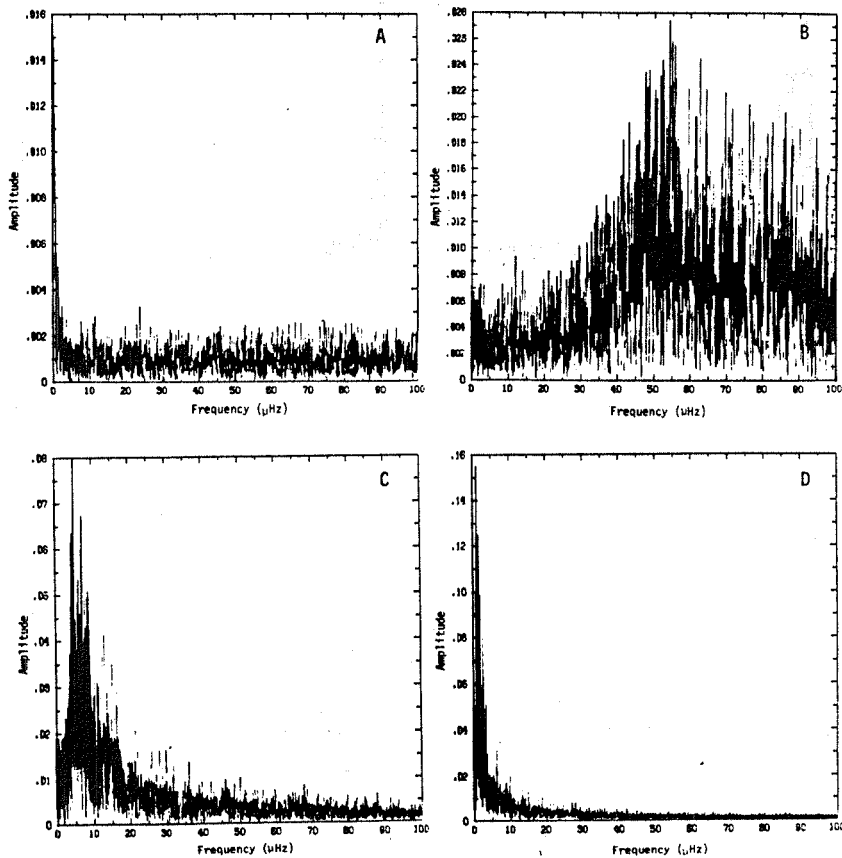
seismology. It should be noted that the 95% duty cycle with 1 hr gaps has its maximum at about 10  $\mu\text{Hz}$ , which might affect searches for rotational splittings of the modes.



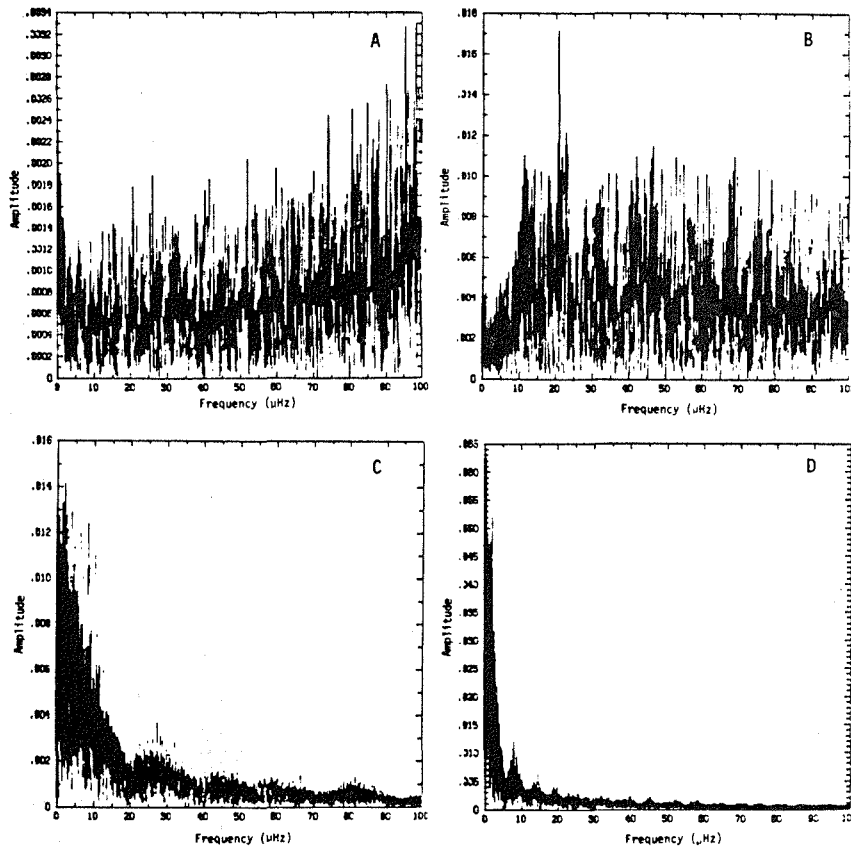
**Figure 2:** Power spectra of window functions resulting from randomly spaced gaps with an average length  $L_g = 60$  min and with varying duty cycles  $D$ . A:  $D=0.80$ . B:  $D=0.85$ . C:  $D=0.90$ . D:  $D=0.95$ . E:  $D=0.97$ . F:  $D=0.99$ .



Figures 3 and 4 show the effect of the average gap length on the window spectrum. Figure 3 illustrates an 80% duty cycle with gap lengths of 6 min (Figure 3A), 60 min (Figure 3B), 600 min (Figure 3C), and 3000 min (Figure 3D). In a space-borne observation context, the short gap length of 6 min might represent data transmission losses, the 600 min gaps might represent minor failures of the TDRS network, and the 3000 min gaps might represent major TDRS failure. Figure 4 illustrates the same gap lengths with a 95% duty cycle. The amplitude of the noise increases as the gap length increases, reaching a maximum of about 10% for the 3000 min case (80% duty cycle). The maximum of the power again varies; for the 95%, 6 min case the maximum occurs at about 138  $\mu$ Hz, which could cause problems when one considers the spacing of the low degree 5 min modes. However, in all cases, it should be noted that the relative power levels of the window functions are quite low, being typically less than 10% of the central power. Thus, probably any of these windows would be suitable for solar oscillation studies.



**Figure 3:** Power spectra of window functions resulting from randomly spaced gaps with a duty cycle  $D = 0.80$  and varying average gap lengths  $L_g$ . A:  $L_g=6$  min. B:  $L_g=60$  min. C:  $L_g=600$  min. D:  $L_g=3000$  min.



**Figure 4:** Power spectra of window functions resulting from randomly spaced gaps with a duty cycle  $D = 0.95$  and varying average gap lengths  $L_g$ . A:  $L_g=6$  min. B:  $L_g=60$  min. C:  $L_g=600$  min. D:  $L_g=3000$  min.

#### IV. Conclusions

It is apparent from these power spectra that almost any of the duty cycles and gap lengths considered here would pose little problem to helioseismological inversions, as long as the gaps are randomly distributed in time. The main goal is to avoid periodic gaps of any sort, as these introduce prominent sidelobes surrounding each solar frequency. It is possible to tailor the window by inserting additional gaps in the data, thereby moving the maximum of the window power spectrum away from regions of interest. However, this of course further degrades the duty cycle, and, if one already has a random distribution of gaps, probably would not be warranted. When one considers the performance of either a satellite-borne instrument, or a ground-based network to obtain solar oscillation data, one must compute the power spectrum of the window function to assess the level of confusion generated by any possible sidelobes.

**Acknowledgements:** It is a pleasure to acknowledge helpful discussions with Juri Toomre, Gordon Newkirk Jr., and Douglas Gough. This work was partially supported by the Air Force Geophysical Laboratory under contracts F19628-77-C-0104 and F19628-83-K-0008 and by NASA through grants NSG-7511 and NAGW-91 to the University of Colorado.

### References

- Bos, R.J., and Hill, H.A.: 1983, Solar Phys., **82**, 89.  
Brown, T.M.: 1979, in Symposium on the Study of the Solar Cycle from Space, NASA CP2098, p.101.  
Claverie, A., Isaak, G.R., McLeod, C.P., Van der Raay, H.B., and Roca Cortes, T.: 1981, Nature, **293**, 443.  
Delache, P., and Scherrer, P.H.: 1983, Nature, **306**, 651.  
Grec, G., Fossat, E., and Pomerantz, M.A.: 1983, Solar Phys., **82**, 55.  
Harvey, J., Pomerantz, M.A., and Duvall, T.: 1982, Sky and Telescope, **64**, 521.  
Stebbins, R., and Wilson, C.: 1983, Solar Phys. **82**, 43.



# ANALYSIS AND INTERPRETATION OF SYNTHETIC TIME STRINGS OF OSCILLATION DATA

Barbara. W. Mihalas  
Jorgen Christensen-Dalsgaard  
Timothy. M. Brown  
High Altitude Observatory  
National Center for Atmospheric Research  
Boulder, CO 80307

**ABSTRACT** Artificial strings of solar oscillation data with gaps and noise, corresponding to the output of different spatial filter functions, are analyzed. Peaks in the power spectrum are identified for values of the degree  $l$  from 0 to 18, and rotational splitting is estimated. The filters prove effective in facilitating identification of essentially all the real peaks in the power spectrum. Estimates of peak frequencies and amplitudes and rotational splitting frequencies are in reasonably good agreement with the input values. "Spurious" peaks in autocorrelation spectra are found to correspond to the frequency spacing between power peaks with the same order  $n$ , differing by one or two in the degree  $l$ .

## I. Introduction

Solar five-minute oscillations divide themselves roughly into three groups, requiring different methods of observation and/or data analysis, according to the value of the angular degree  $l$ . As  $l$  increases for a fixed value of the order  $n$ , modes differing by one in  $l$  become increasingly densely packed in frequency  $\nu$ ; modes differing only in  $n$  lie on distinct and clearly separated ridges in the  $(l - \nu)$  plane (see e.g. Deubner 1975; Rhodes et al 1977). The large amount of power in the p mode ridges at  $l$ -values greater than about 200 is the result of the high density of modes, each mode having a velocity amplitude of a few  $\text{cm s}^{-1}$ ; in this regime there is no hope of seeing individual modes with distinct  $l$ -values. The high- $l$  modes are detected by making spatially resolved time series observations, usually of Doppler shifts. A Fourier transform in both space and time then produces (from good data) the familiar ridge structure, if the time series is long enough and the analysis is done carefully.

At the other extreme,  $l=0-3$ , the five minute modes have  $n \gg l$  and the frequencies are given by

$$\nu_{nl} \approx \Delta\nu (n+l/2+\epsilon) + \text{small term in } l(l+1) \quad (1)$$

where  $\Delta\nu$  is of the order of 140  $\mu\text{Hz}$  and  $\epsilon$  is a fixed offset (see Vandakurov 1967). Thus the modes are distinct in both  $n$  and  $l$ . The frequency difference between modes with successive values of  $n$  and the same  $l$  is  $\Delta\nu$ , while that between successive  $l$ -values is half as large,  $\Delta\nu/2$ . Observations of at least a few hours duration (needed to obtain an adequate signal-to-noise ratio) in integrated light from a large part or all of the solar disk yield only these very low- $l$  modes (Grec et al 1980, 1983); these observations average upward and downward velocities (or positive and negative intensity fluctuations) over the entire sun, thus strongly reducing the signals from modes with angular degree

greater than about 3 or 4 and in that sense acting as a spatial filter. Whole-disk measurements have the advantage that for the very low  $l$  modes spatial averaging increases the signal-to-noise ratio.

Taking the temporal Fourier transform of a whole-disk velocity signal yields a very regular pattern in frequency. Peaks appear evenly spaced at intervals of  $\Delta\nu/2$ , like a picket fence. Alternate peaks have the same  $l$ -value, so the spectrum represents one set of pickets for modes with  $l=0$  and 2 at intervals  $\Delta\nu$ , interleaved (at separation  $\Delta\nu/2$ ) with a second set for  $l=1$  and 3. The second term in equation (1) produces a very small separation between  $l=0$  and  $l=2$  peaks, and between  $l=1$  and  $l=3$ .

Because the pattern is so regular, each segment of the power spectrum of length  $\Delta\nu$  in frequency looks nearly identical to the next segment of length  $\Delta\nu$ , in terms of peak positions (the amplitudes may vary). If two successive segments are superimposed, all the peaks overlay each other almost perfectly; if summed, the peaks stand out more clearly against the noise, which is randomly distributed. Grec et al (1980) found that by cutting the whole power spectrum into lengths  $\Delta\nu$  and summing to produce a "superposition spectrum" of total length  $\Delta\nu$ , the peaks for  $l=0-3$  became much more distinct.

Modes with  $4 < l < 200$  are not sufficiently densely packed in frequency to produce much power in the p mode ridges, and they have too rapid variation in velocity across the solar disk to be seen in integrated sunlight. They have only recently been seen with any clarity in spatially resolved observations as they require extremely long duration time series (Duvall and Harvey 1983).

In this paper we report on an analysis, using artificial data, designed to answer some important questions about modes with  $l$ -values from 0 to 18 (many of the results can be readily extrapolated to somewhat higher  $l$ -values). There has been considerable interest recently in the question of whether spatial filters can be constructed that will isolate sets of modes for  $l \geq 3$  in such a way that the intermediate- $l$  modes can be unambiguously identified from frequency power spectra. The procedure to be discussed presupposes spatially resolved observations, in two spatial dimensions. The Fourier Tachometer (see Brown, these proceedings) is ideally suited for making the required measurements.

In principle, separation of modes with different spatial wavenumbers is accomplished by taking a spatial Fourier transform; for each particular choice of the wavenumber  $\mathbf{k}_0$ , the transform  $G(\mathbf{k}_0, t)$  should have no contamination from any other value of  $\mathbf{k}$ . In fact, the functions that describe spherical modes are orthogonal only when integrated over the entire sphere, while observations are made over less than half, when limb effects are taken into account. As a result, the spatial transform function in spherical geometry  $F_{l_0}(\vartheta, \phi)$ , optimized for degree  $l_0$  but integrated only over the visible part of the Sun and operating on the observed Doppler velocity  $v(\vartheta, \phi, t)$ , yields a function

$$G_{l_0}(l_0, l_0 \pm 1, l_0 \pm 2; t) = \int v(\vartheta, \phi, t) F_{l_0}(\vartheta, \phi) d\Omega \quad (2)$$

which at best contains strong components of these five  $l$  values and very weak components of many others. Here  $d\Omega$  indicates an element of solid angle; elsewhere  $\Omega$  will always denote angular rotation frequency.

The spatial function  $F_l$  is often called a "spatial filter", as it filters out modes with all but a few  $l$  values. Integrating the data over a succession of filters, each centered on a different  $l_0$ , produces a series of time strings, each of which can be Fourier analyzed in time to produce a frequency power spectrum containing peaks with significant amplitudes for only five  $l$  values. The difficulty that arises in analyzing such spectra is that, starting with modes of degree  $l=3$  or 4, the frequency separation between a mode with a particular  $(l, n)$  and that with  $(l, n+1)$  approximately equals the frequency separation between  $(l, n)$  and  $(l+3, n)$ , so that the peaks for  $(l, n+1)$  and  $(l+3, n)$  are nearly coincident in the power versus frequency spectrum.

To examine the question of isolating and identifying modes by using spatial filters, we have engaged in an exercise dubbed "Hide' and Seek" by D. O. Gough, in which one of us constructs a set of filtered time series and presents them to another of us, providing only the information that an observer would have about his data: duration, sampling interval, duty cycle, and the relative amount by which each filter should reduce the amplitude of each mode. The second player deduces from the data as much information as possible in a reasonable amount of time and presents the information to the omniscient first player who compares it with the input data. The data analyzer does not, of course, discuss the analysis with the data creator until ready to relinquish the analysis. Similar hide and seek experiments have been done by two of us (TMB + JC-D) to study the effectiveness of T. Brown's maximum entropy method gap-filling routine.

In § II, we will first describe how the data sets were created and how the power spectra were produced. Then in § III we will recount chronologically the few major snags encountered in trying to understand the power spectra; these snags proved to be quite informative. The methods used to obtain the frequencies, frequency splittings and amplitudes are then described.

Section IV is a summary of the main results obtained from the analysis and a comparison of these results with the input data, plus a brief discussion. A number of ideas for future experiments are presented in § V.

## II. Time Series Construction

The time series of (filtered) artificial data were constructed by JC-D. The velocity of each mode was assumed to be given by

$$v_{n,l,m}(t) = \hat{V}_{n,l,m} \exp[-0.25(\nu - \nu_\mu)^2 / \nu_0^2] \sin(2\pi\nu_{nlm}t + \varepsilon_{nlm}) \quad (3)$$

where  $\hat{V}_{nlm}$  is a time-independent amplitude,  $\nu_\mu = 3.18\mu\text{Hz}$ , and  $\nu_0 = 0.35\mu\text{Hz}$ . The squared amplitudes  $\hat{V}_{nlm}^2$  were exponentially (and randomly) distributed (see Christensen-Dalsgaard and Gough, 1982), with  $\text{rms}(\hat{V}_{nlm}) = 15 \text{ cm } \sim\text{s}^{-1}$ . The  $\varepsilon_{nlm}$  are time-independent random phases.

The velocities  $\{v_{nlm}(t)\}$  used for the time series are defined in wavenumber space, while velocities in real data are functions of position (and not wavenumber), e.g.  $v(\mathbf{x}, t)$  or  $v(\vartheta, \phi, t)$ . To create a velocity time series, whose  $l$  dependence is centered on  $l_0$ , i.e. an artificial time string equivalent to  $G_{l_0}$  in

equation (2), the  $\{v_{nlm}\}$  velocities are weighted by filter functions  $S_{lm}(l_0)$  in wavenumber space and summed. An artificial velocity time string is then given by

$$V(l_0, t) = \sum_{l, m, n} S_{lm}(l_0) v_{nlm}(t) + N(t) \quad (4)$$

where  $N(t)$  is random Gaussian noise with rms amplitude 20 cm s<sup>-1</sup>. The "wavenumber filter functions"  $\{S_{lm}(l_0)\}$  operating on  $\{v_{nlm}(t)\}$  are designed to produce the same response as the spatial filter functions  $\{F_{l_0}(\vartheta, \phi)\}$  operating on real velocities  $v(\vartheta, \phi, t)$ .

The rotationally split frequencies are given by

$$\nu_{n,l,m} = \nu_{n,l,0} + m \Delta\nu_{nl}^{(1)} + m^2 \Delta\nu_{nl}^{(2)} \quad (5)$$

where  $\nu_{n,l,0}$  is taken from model 1 of Christensen-Dalsgaard (1982) and the rotational splitting corrections are

$$\Delta\nu_{nl}^{(1)} = \frac{\beta_{nl}}{2\pi} \int_0^1 K_{nl}(r) \Omega(r) dr \quad (6a)$$

$$\Delta\nu_{nl}^{(2)} = 4000 \frac{\nu_{nl,0}}{n} \frac{1}{8l(l+1)-6} [\Omega_s^2 R/g_s + 3J_2] \quad (6b)$$

(Goode, these proceedings). Here  $K_{nl}$  is the rotational splitting kernel for mode  $(n, l)$ ,  $\Omega(r)$  is the rotation rate as a function of radius  $r$ ,  $\nu_{nl,0}$  is the unperturbed mode frequency,  $\Omega_s$  the surface rotation rate,  $R$  the solar radius,  $g_s$  the surface gravity,  $J_2$  the second moment of inertia and  $\beta_{nl}$  is a constant of order unity. The rotation curve  $\Omega(r)$  used here was the same curve that is discussed in Christensen-Dalsgaard and Gough in these proceedings. Frequencies  $\nu_{n,l,0}$  in the range  $2.5 \leq \nu_{nl} \leq 3.8$  mHz and  $0 \leq l \leq 18$  were included.

The "observer", BWM, had none of the foregoing information. The "observing data" given her was: sampling interval = 2 min.; "observing sequence" for data: 11 strings of 18.8 hour "runs" followed by 5.2 hour gaps, for a duty cycle of about 0.8 and total observing time of about 207 hours; filter values for each  $l$ , for each time series (for typical examples of an optimized filter, see JC-D, 1984, Figure 9, these proceedings). The sampling interval and total observing time were limited by the total number of data points we could process on the computer we were using (HAO's PDP 11/70).

To obtain power spectra, the time series were passed first through T. Brown's gap-filling routine, which uses a maximum entropy method. The resulting continuous time series were then Fourier analyzed in time. The output consisted of 18 temporal power spectra covering the frequency range 2.1-4.4 mHz, with frequency resolution about 1  $\mu$ Hz. (which is the theoretical resolution of the  $\sim 210$  hr time string. Although we plan to implement frequency oversampling in future experiments, computer memory size limitations made that impossible for the analysis reported here.)



### III. Procedure for Analyzing Data

**Step 1:** The obvious way to begin analyzing power spectra known to have just a few  $l$ -values present is to use the tactics developed for disk-averaged ( $l=0-3$ ) data. These include (a) visual examination of the frequency power spectrum; (b) taking the autocorrelation of a power spectrum as a function of frequency shift. This procedure tends to produce large values of the autocorrelation function at frequency shifts equal to the spacing interval of a series of power peaks evenly spaced in frequency; (c) superposition of segments of the power spectrum as described in the introduction (see also Grec et.al. 1980, Fig.2). If the segment length is chosen "correctly" it will equal  $\nu_{n+1,l} - \nu_{n,l}$ , which, for a given  $l$  and for  $l \ll n$ , is almost independent of  $n$  (Vandakurov 1967).

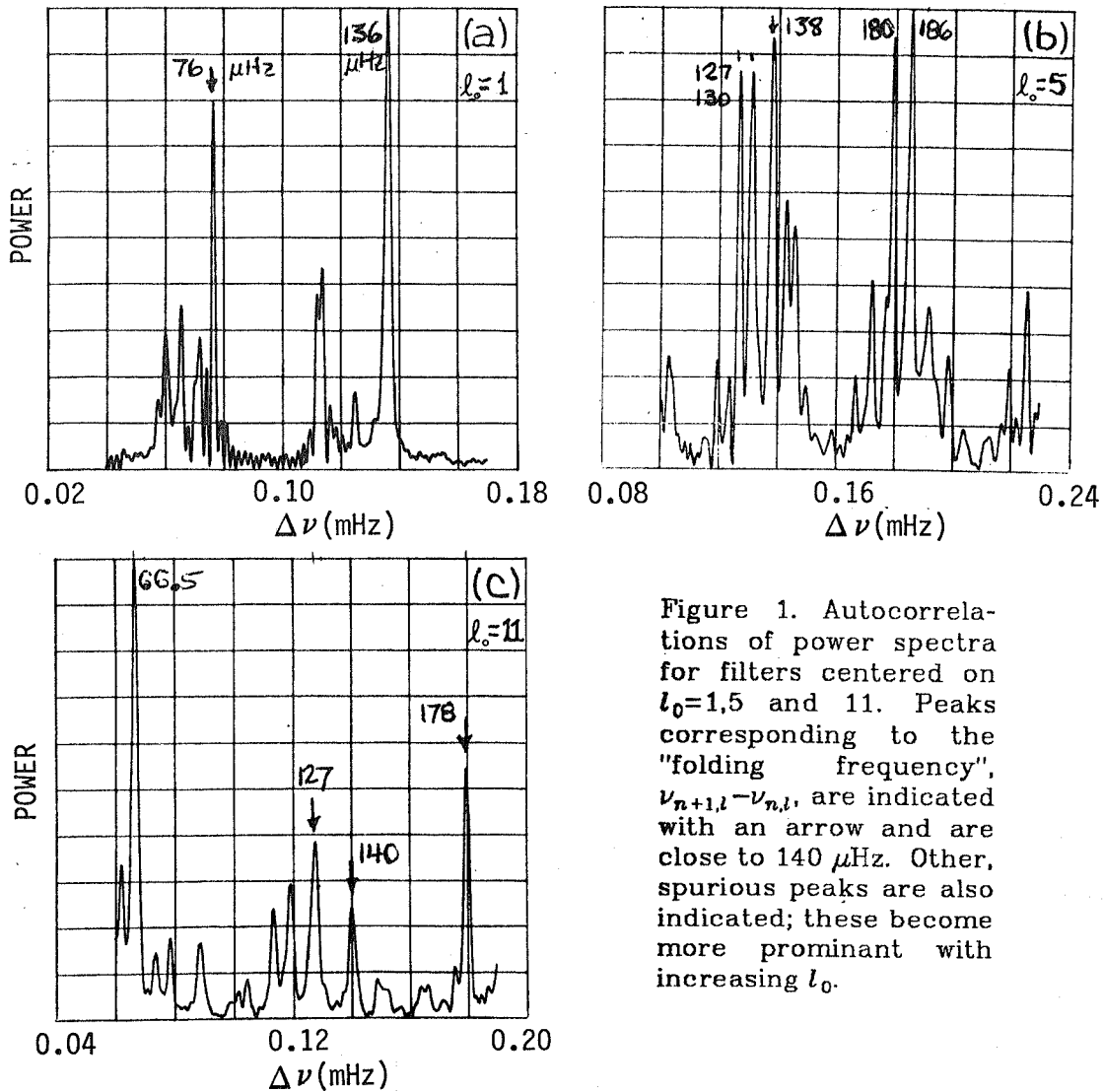


Figure 1. Autocorrelations of power spectra for filters centered on  $l_0=1,5$  and  $11$ . Peaks corresponding to the "folding frequency",  $\nu_{n+1,l} - \nu_{n,l}$ , are indicated with an arrow and are close to  $140 \mu\text{Hz}$ . Other, spurious peaks are also indicated; these become more prominent with increasing  $l_0$ .

We assigned to each filter the number of the degree  $l_0$  on which it was centered; thus filters 1-18 were centered respectively on  $l_0 = 1-18$ . The first few filters, which isolate  $l$  values of zero to about 5, have autocorrelation functions

that exhibit a dominant peak at roughly 140mHz (more precisely, 136-138 $\mu$ Hz). By filter 5, however, which isolates  $l = 3-7$ , other peaks in the autocorrelation function are stronger than the 140 $\mu$ Hz peak, an effect which becomes increasingly pronounced with an increase in the  $l$  value for which the filter was optimized. Figure 1, showing the autocorrelation functions for filters 1, 5 and 11, clearly exhibits this effect.

A similar difficulty arose with the superposition spectra (defined in the introduction). For each filter, the superposition length  $\Delta\nu(l_0)$  was taken to be the center of the peak nearest 140 $\mu$ Hz in the autocorrelation spectrum. The power spectrum was then cut into pieces of length,  $\Delta\nu(l_0)$ , which were summed. The resulting power spectrum, for  $l = 0-3$ , shows four peaks, one for each  $l$ -value (with a very small amplitude for the  $l = 3$  peak). However, for the filter centered on e.g.  $l_0 = 5$ , the superposed spectrum is virtually impossible to interpret without further knowledge. Examples for filters centered on  $l_0 = 1, 5$  and 11, shown in Figure 2, illustrate the problem.

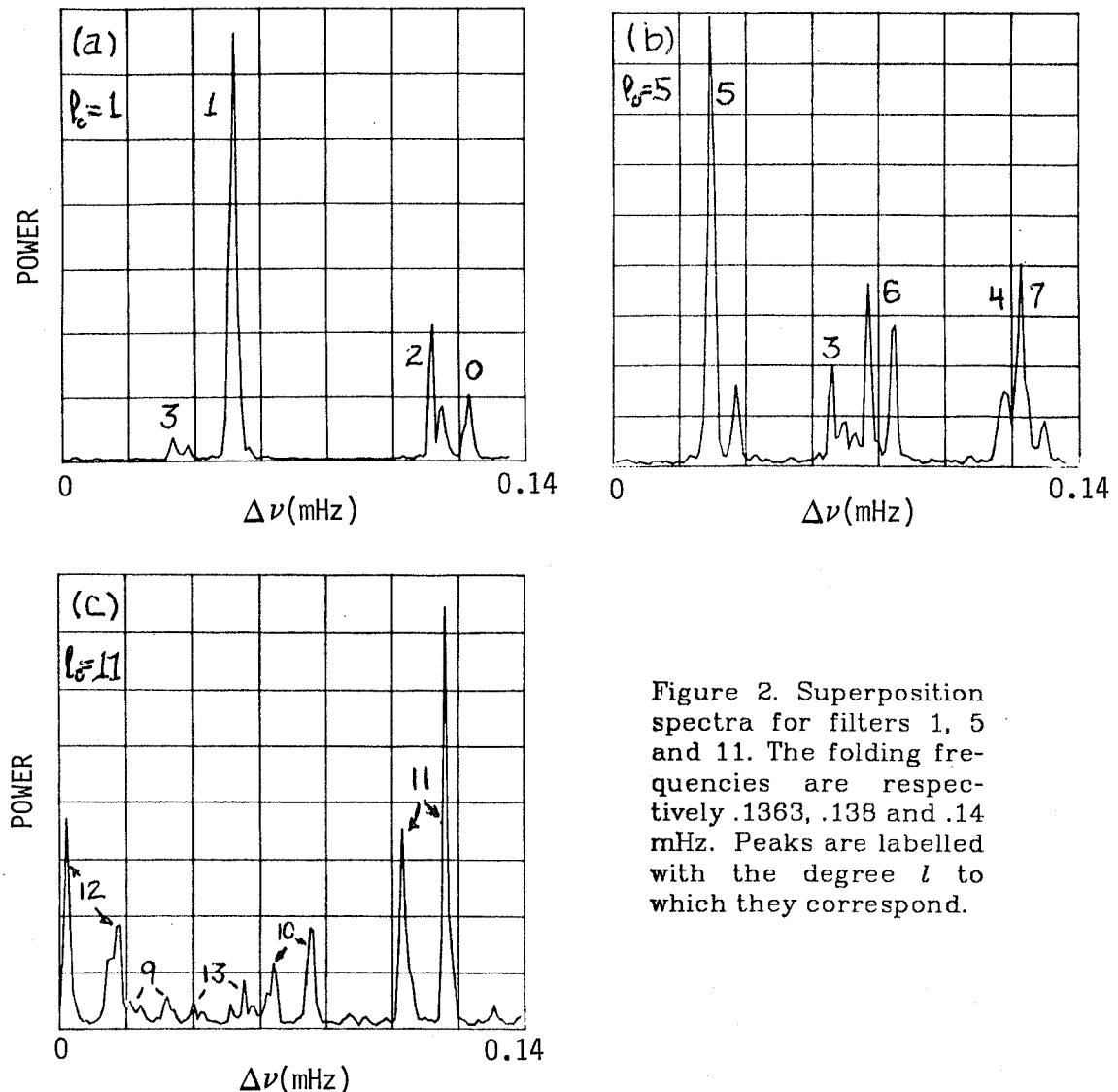


Figure 2. Superposition spectra for filters 1, 5 and 11. The folding frequencies are respectively .1363, .138 and .14 mHz. Peaks are labelled with the degree  $l$  to which they correspond.

**Step 2:** Because in fact the segments of the superposition spectrum vary slightly from one to the next, it can be more informative to lay them one below another like rungs on a ladder, in what is called an "echelle diagram." For folding frequency  $\Delta\nu(l_0)$  each successive layer begins at a frequency  $\Delta\nu(l_0)$  greater than the beginning of the one above it, as shown on the left-hand side of Figure 3c, which has a folding frequency of almost exactly 140  $\mu\text{Hz}$ . For each segment, the power was normalized by its maximum value within that segment, in order to make the peaks more visible. We note that this normalization exhibits clearly the frequency ranges in which no clear peaks emerge above the noise.

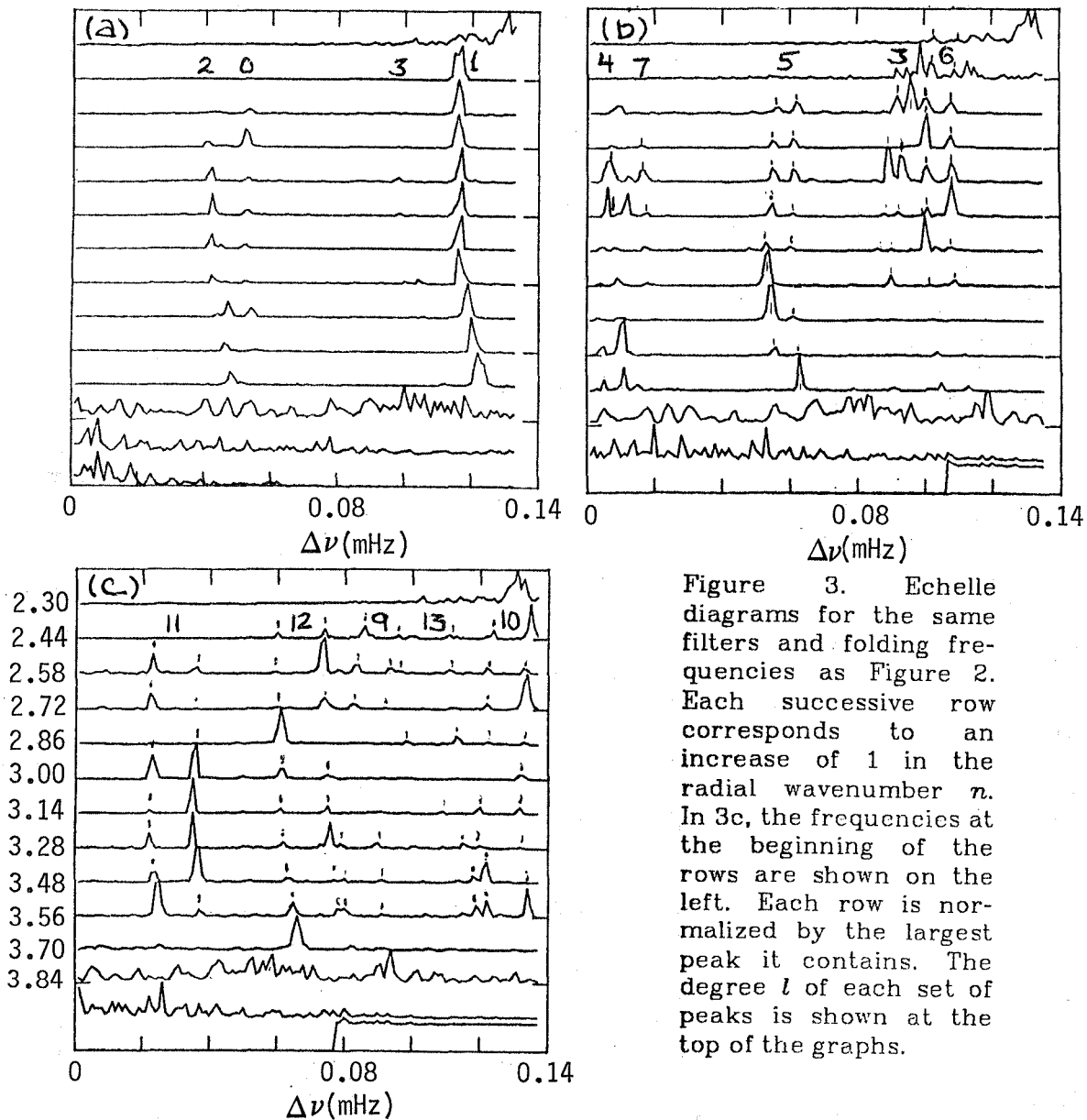


Figure 3. Echelle diagrams for the same filters and folding frequencies as Figure 2. Each successive row corresponds to an increase of 1 in the radial wavenumber  $n$ . In 3c, the frequencies at the beginning of the rows are shown on the left. Each row is normalized by the largest peak it contains. The degree  $l$  of each set of peaks is shown at the top of the graphs.

We checked to see whether such echelle diagrams (see e.g. Grec et al. 1983, Fig. 9 for an echelle diagram which plots frequency only) would be easier to interpret than superpositions. Figure 3, which compares echelle diagrams for filters 1, 5 and 11, demonstrates that this approach also fails to produce clearly identifiable peaks for some filters; detailed scrutiny establishes that for the filters and frequencies where confusion reigns in the superpositions, little improvement is achieved with the echelle plots.

The echelle plots do show clearly, however, for sets of isolated peaks, that the frequency difference  $\Delta\nu_{n,n+1}(l)$  is a weak function of  $n$ ; if it were constant with  $n$ , all peaks on one echelle plot would line up exactly vertically or (if  $\Delta\nu_{n,n+1}$  was wrong for the chosen  $l$  value) would fall on a straight diagonal line. Instead, they fall on a nonlinear curve. The sequence of echelle plots shows the decrease in  $\Delta\nu_{l,l+2}(n)$  with increasing  $l$ .

Perhaps the most useful information in the echelle plots concerns the question of rotational splitting. If each mode  $(l,n)$  produces two peaks (with  $l = \pm m$ ) that survive the filter and that are separated in frequency by  $\Delta\nu_{l,-l} = 2l\Delta\nu_{m,m+1}$ , then a superposition of all peaks for a given  $l_0$ , and various  $n$  values, should be made in such a way that the unperturbed frequencies  $\{\nu_n(l_0), n = n_1, \dots\}$  all coincide. Such superposition requires some care, as  $\Delta\nu_{n,n+1}(l_0)$  is not necessarily a simple function of  $n$ . We note also that  $\Delta\nu_{n,n+1}(l)$  is quite a strong function of  $l$ , hence its mean value changes significantly with  $l$ .

*Step 3:* Ultimately, it was necessary to use the information that  $\Delta\nu_{n,n+1}$  (fixed  $l$ ) and  $\Delta\nu_{l,l+1}$  (fixed  $n$ ) change continuously and gradually with  $l$  and  $n$ . This must be so because the eigenfrequencies are determined by the structure of the solar model, and are most affected by regions of the model where the corresponding eigenfunction is large. But in the 5 minute band, eigenfunctions change little from  $(n,l)$  to  $(n+1,l)$  or from  $(n,l)$  to  $(n,l+1)$ .

Frequencies for  $l = 0, 1, 2, 3$  modes can be determined unambiguously from echelle diagrams of filter 1, 2, 3 power spectra. The frequency spacings  $\Delta\nu_{n,n+1}(l = 3)$  and  $\Delta\nu_{l,l+1}(n = 3)$  can be used to determine approximate locations for  $\{\nu_{n,l=4}\}$  peaks. A search of the immediate frequency neighborhood for peaks, combined with the constraint that variations in  $\Delta\nu_{n,n+1;l_0}$  and  $\Delta\nu_{n;l,l+1}$  be reasonably smooth, proved sufficient to locate full sets of likely looking peaks for  $l = 4$ . The frequency separations  $\Delta\nu_{n,n+1;l=4}$  and  $\Delta\nu_{n;l=3,l=4}$ , combined with the trends that appeared in the mean separations  $\Delta\nu_{n,n+1;l}$  and  $\Delta\nu_{n;l,l+1}$ , both for fixed  $l$ , then provided an estimate of the positions for  $\{\nu_{n,l=5}, n = \dots\}$ .

For each filter, peaks for  $l_0$  and for  $l_0 \pm 1$  and  $l_0 \pm 2$  would be visible. The range in which peaks were difficult to separate was that for which  $\Delta\nu_{n,n+1;l} \approx \Delta\nu_{n;l,l+3}$  and occurred mainly for  $l$  values in the range  $l = 4 - 10$ . Thus  $l_0 - 2$  and  $l_0 + 1$  would have peaks that nearly coincided in frequency, as would  $l_0 - 1$  and  $l_0 + 2$ . The peaks for  $l_0$  modes, however, had almost no contamination, because the amplitudes of  $l_0 \pm 3$  peaks were diminished by the filter to less than 5% of their unfiltered value, while those of  $l = l_0$  modes were greater than 50% of their unfiltered value. Each filter was used primarily for modes of its particular  $l_0$ , stepping through the filters one filter and thus one  $l_0$  at a time.

The peak frequencies were checked for consistency by plotting  $\Delta\nu_{n,n+1}$  and  $\Delta\nu_{l,l+1}$  against  $l$ , as shown in Figures 4a and 4b. A further check was provided by making a standard ( $l-\nu$ ) diagnostic diagram, shown in Figure 5.

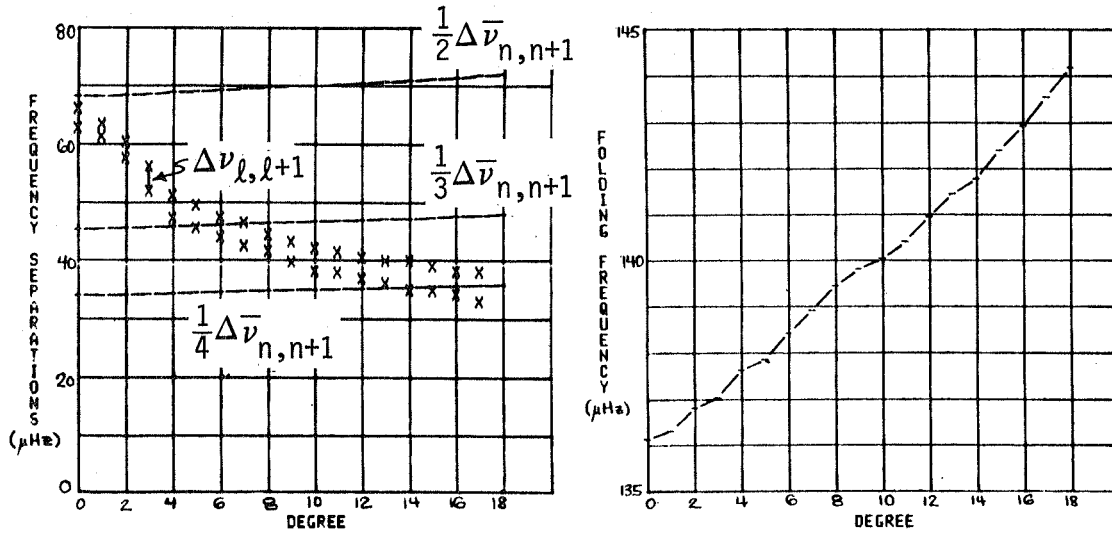


Figure 4a. Folding frequency  $\nu_{n+1,l} - \nu_{n,l}$  versus  $l$ , in  $\mu\text{Hz}$ . These are mean values, averaged over  $n$  for each  $l$ .

Figure 4b. Frequency spacings, in  $\mu\text{Hz}$ , versus  $l$ . For each  $l$ , the range of values as  $n$  varies is shown of  $\Delta\nu_{l,l+1} \equiv \nu_{l+1} - \nu_l$ ; for each  $l$ , the minimum and maximum values of  $\Delta\nu_{l,l+1}$  are indicated by 'x'. For comparison,  $1/2$ ,  $1/3$  and  $1/4$  of the folding frequency, averaged over  $n$ , are shown. As  $l$  increases, the folding frequency goes from roughly twice to three times to four times the frequency spacing in  $l$ .

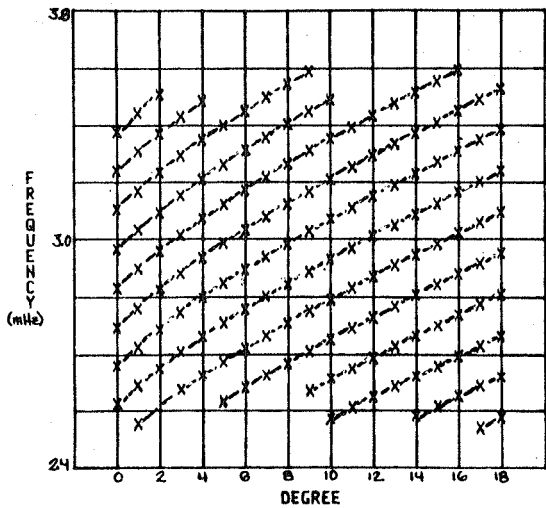


Figure 5. Diagnostic diagram of  $\nu_{n,l}$  versus  $l$ . Dashed lines connect frequencies with the same order  $n$ .

Figure 4b also displays clearly why a problem arises with near-coincidence between peaks of different  $l$ -values. The separation between peaks with successive  $l$ -values and the same radial wavenumber  $n$  decreases with increasing  $l$ , while the frequency separation between peaks with the same  $l$  and successive  $n$ -values increases. For  $l=0-2$ ,  $\Delta\nu_{n,n+1}(\text{fixed } l)$  is slightly larger than  $\Delta\nu_{l,l+3}(\text{fixed } n)$ ; the difference is great enough that the  $(n, l+2)$  peak does not lie on top of the  $(n+1, l)$  peak. For  $l=4-7$ , however,  $\Delta\nu_{n,n+1} \approx \Delta\nu_{l,l+3}(\text{fixed } n)$  and peaks differing by 3 in  $l$  and by one in  $n$  have nearly the same frequency. A similar problem arises starting at about  $l=13$  between peaks with  $(n, l+4)$  and those with  $(n+1, l)$ ; the filtered time strings are not strongly affected by this coincidence, however, because for a filter centered on  $l_0$ , only the  $l_0 \pm 2$  peaks are involved.

In Figure 4b we also see the reason for "spurious" peaks in the autocorrelation spectra. In the low- $l$  cases there were just two regular intervals between peaks, the folding frequency  $\Delta\nu_{n,n+1}(l)$  and half the folding frequency. At higher  $l$ -values  $\Delta\nu_{l,l+1}$  and  $\Delta\nu_{n,n+1}$  are no longer nearly commensurate. Thus there are also series of peaks regularly spaced at intervals  $\Delta\nu_{l,l+1}(n)$ ,  $\Delta\nu_{l,l+2}(n)$  and  $\Delta\nu_{l,l+3}(n)$ , which may all differ significantly from the folding frequency  $\Delta\nu_{n,n+1}(l)$ .

**Step 4:** In step 3 the frequencies  $\{\nu_{n,l}\}$  (for  $l = 0 - 19$  and as many  $n$  values as were reasonably unambiguous) were found to the nearest frequency point in the power spectrum. The next step was to develop better approximations to the frequencies by either (a) fitting the central 3 points of the peak with a parabola and determining the position of its maximum or; (b) doing a centroid on the central 3 points or the central 5 points of the peak. In case (a), the power in the peak was estimated by integrating the area between the two frequencies where the parabola crossed the  $\nu$ -axis; in case (b), it was given by the sum of the power at the central 3 (5) points. The amplitude of a mode is obtained (approximately) by taking the square root of the power and then dividing by the filter function.

Amplitudes estimated by parabola fitting and from 3 and 5 point centroids are shown for a typical case in Figure 6; for all the  $n$  values found for a particular  $l$  (here  $l=3$ ), the estimates are shown for the peak amplitudes in the power spectra of different filters. The three methods tend to produce estimates that retain roughly the same ratio from peak to peak.

Finally, rotational splitting was calculated for each mode  $(l, n)$  as

$$|\Delta\nu_{m,m+1}(l, n)| \approx (2l)^{-1} |\nu_{l,n,m=+l} - \nu_{l,n,m=-l}|. \quad (7)$$

These were calculated separately for parabola fitting and for 3-point and 5-point centroids. For  $l = 1, 2$  the measurements were difficult and highly inaccurate, but the accuracy improved rapidly as  $l$  increased. To further improve accuracy, the average over  $n$  of  $\Delta\nu_{m,m+1}(l, n)$  was taken for each  $l$ . This is shown on Figure 7, along with the "correct" curve obtained from the frequencies  $\nu_{nlm}$  defined by equations (5)-(7), which were used as input to the velocity times strings [see equation (3)].

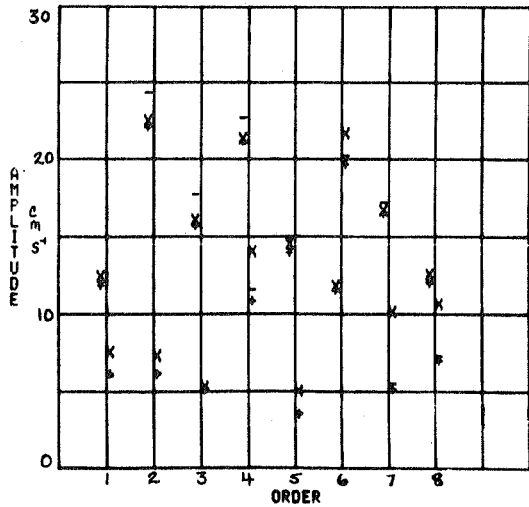


Figure 6. Amplitudes of  $l=2$  modes, estimated from the power spectra from filter 3. The '-', '+' and 'x' points are from parabola, 3-point centroid and 5-point centroid estimates respectively. The order numbering starts with the smallest-frequency peak in the data; its actual  $n$ -value could not be determined from the data. Amplitudes of  $l=\pm m$  peaks of mode  $(l,n)$  are shown just to the right and left of the vertical at  $n$ . The abscissa also represents frequency, with  $\nu$  increasing by  $\sim 137 \mu\text{Hz}$  when  $n$  increases by 1.

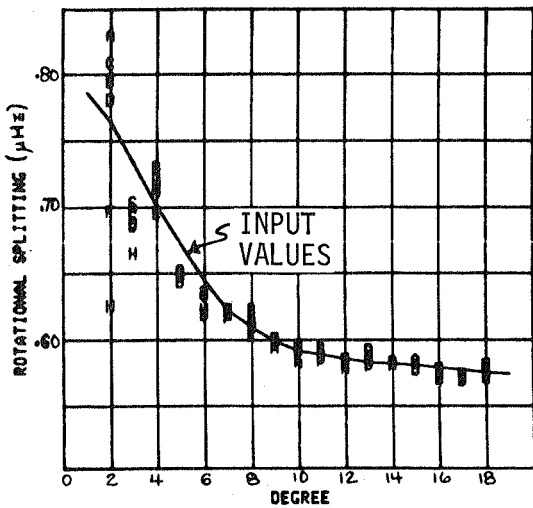


Figure 7. Rotational splitting,  $(\nu_{n,l,m=+l} - \nu_{n,l,m=-l}) / 2l$ , averaged over all  $n$  values for each  $l$  and plotted against  $l$ . The solid curve was calculated from the input rotation curve  $\Omega(r)$ . Points labelled 'A,B,C,D,F,H' correspond to values obtained from parabola fitting, straight (A) and weighted (B) averages, 3-point centroid (C and D) and 5-point centroid (F and H) fits to peaks in the data.

#### IV. Results and Discussion

It turned out that both of the rotationally split components of nearly all the input modes were detectable, down to input amplitudes of about  $1\text{-}2 \text{ cm s}^{-1}$ . Exceptions were all at the high and low frequency ends of the 5-minute band. This demonstrates that the intermediate- $l$  modes can be isolated and identified by the use of optimized filters, without having to resort to viewing the sun from two widely separated (optimally  $90^\circ$ ) angles.

The frequencies and amplitudes derived from the data by BWM were compared by JC-D with the input values. RMS deviations in the frequencies were of the order of  $0.5 \mu\text{Hz}$ , and were little better when only large amplitude modes were included. The frequency differences showed some systematic dependence on frequency which can be seen in Figure 8a. This systematic dependence may result from a type of interference between the gaps and the frequencies; a

similar "hide and seek" experiment with gapless data will demonstrate whether the gaps are the culprit. Dependence on amplitude is shown in Figure 8b, and shows no particular systematic trends.

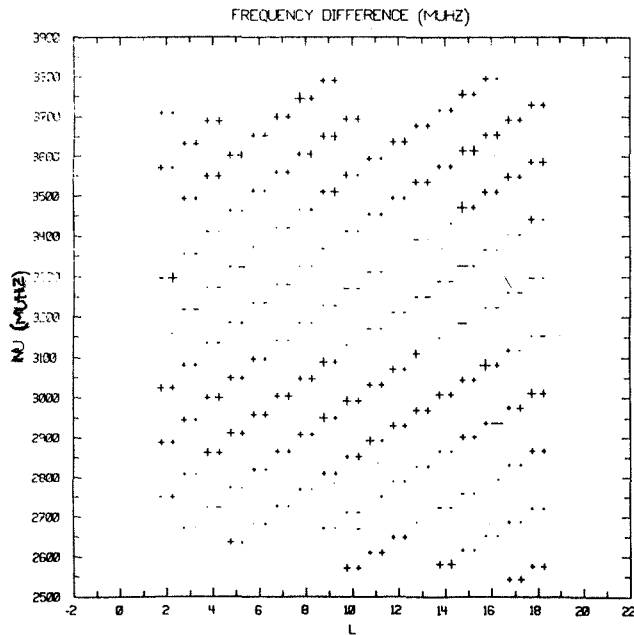


Figure 8a. Difference between input frequency and frequency estimated from the artificial data sets, as a function of degree  $L$ . The symbols '+' and '-' indicate the sign of the difference; size of the symbol indicates the magnitude. At frequencies of about 2.85-3.1 and 3.5-3.8 mHz the differences are almost uniformly positive; below 2.85 and from 3.1-3.5 mHz they are like bands of negative differences.

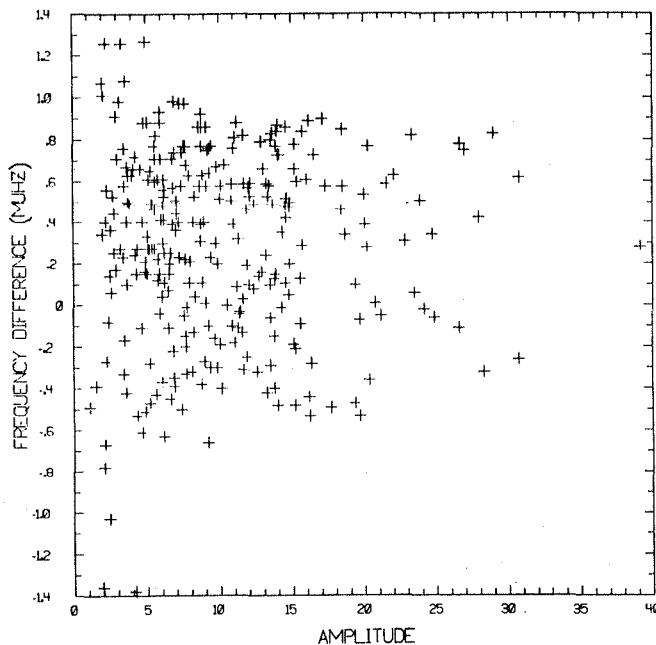


Figure 8b. Same as Figure 8a, except that the frequency differences are plotted against amplitude, and the sign of the '+' marks has no significance. Here no systematic effects like those in Figure 8a are seen.

Amplitude ratios are shown in Figure 9 as a function of mode amplitude. Although the accuracy of amplitude estimates is poor compared to that of frequency estimates we note that almost all inferred amplitudes are within a factor of two of being "correct". This is several orders of magnitude better than our theoretical uncertainties in mode energies and thus indicates that amplitudes derived from Doppler velocity signals may be useful.



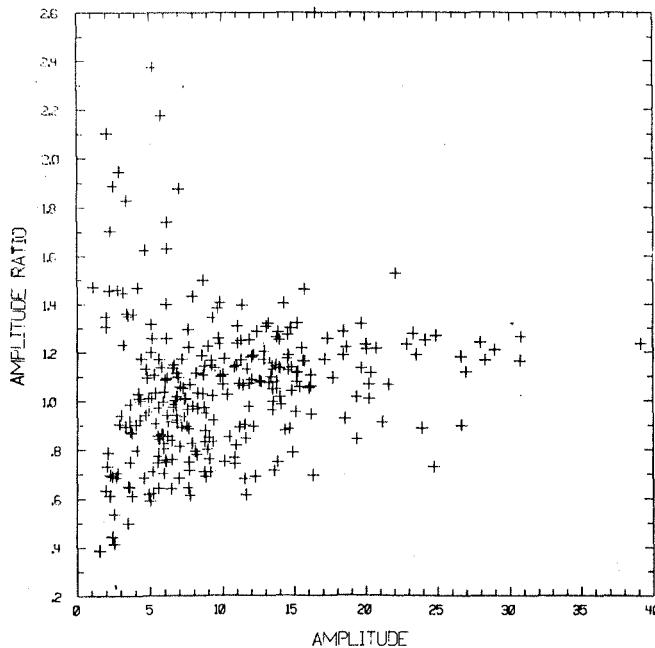


Figure 9. Ratio of estimated amplitude to input amplitude plotted against amplitude. There appears a slight tendency to overestimate the amplitude from the data. As expected, the scatter is greatest for very small amplitude peaks.

Figure 7 shows the mean rotational splitting for the input data compared with the mean values obtained from parabola fitting and 3 and 5 point centroids. The expected increase in accuracy with increasing  $l$  is clearly seen.

Inaccuracies in the frequencies (and thus also in the rotation curve) appear to result mainly from three effects. One is the limit of about  $1 \mu\text{Hz}$  on the frequency resolution in the power spectrum. Precision can probably be improved by oversampling in frequency by a factor of 4 or 5; such experiments will be tried as soon as the data reduction codes are running on a larger computer.

The second is the presence of noise, which probably has its effect mainly on small-amplitude modes. One can picture the effect readily by considering a peak located halfway between two pixels, with negligible amplitude on the other neighboring pixels. Adding noise at an amplitude  $1/4$  that of the "real" signal on one pixel will change the apparent frequency by about  $.05 \mu\text{Hz}$ , while noise with the same amplitude as the peak shifts the frequency by  $.17 \mu\text{Hz}$ .

The third is the intrinsic limitation imposed by the finite length of the data string and by the gaps in the data. In real data finite mode lifetimes provide another source of uncertainty in frequency.

## V. Future Experiments

There are numerous other experiments of this general sort that can help elucidate the relationship between what the Sun puts into pulsation modes and what terrestrial observers deduce out of their data:

- (a) An analysis identical to that reported here can be done with 60% and 100% duty cycles to estimate the frequency uncertainty caused by data gaps and by gap filling. Similarly we should check how much improvement is obtained

in the frequencies by taking data strings 2-3 times as long.

- (b) Identical analysis should also be done on identical time series with an over-sampled power spectrum.
- (c) Finite mode lifetimes can be synthesized by randomly varying the amplitude and phase of the velocity signal at time intervals corresponding to some assumed lifetime. The effect on apparent frequencies and amplitudes can be deduced from the power spectra.
- (d) Experiments with more realistic noise models should be carried out. Attempts should be made to model terrestrial ("seeing") noise and the solar noise arising from granules and supergranules. These experiments should also be combined with (a).
- (e) Legendre polynomial filters should be tried in place of the optimized Chebyshev filters; Legendre polynomials are the natural filter function for normal modes of a sphere. These could also be used to isolate a range of  $m$  values other than  $m = \pm l$ .
- (f) Finally, the circle of deductive-inductive analysis should be closed by using the deduced frequencies and rotational splittings as input for model inversion and rotational inversion calculations (e.g., Christensen-Dalsgaard and Gough 1984, these proceedings). This would necessitate creating artificial data strings for higher- $l$   $p$  modes, for  $g$  modes and probably for long-period  $p$  modes, and performing on these data the sort of analysis reported here to obtain frequencies and rotational splitting. One could then estimate how badly the structure of the Sun itself is masked by effects of noise, both solar and terrestrial, and by finite data strings, finite mode lifetimes and data gaps. This, of course, is the bottom line for all of our experiments, observations and analysis.

#### Acknowledgements

We wish to thank Dr. P. A. Gilman for reviewing the manuscript and for making some useful suggestions. The National Center for Atmospheric Research is sponsored by the National Science Foundation.

#### References

- Christensen-Dalsgaard, J. 1982, *M.N.R.A.S.* 199, 735
- Christensen-Dalsgaard, J. and Gough, D.O., 1982, *M.N.R.A.S.* 198, 141
- Deubner, F.-L. 1975, *Astron. Astrophys.* 44, 371
- Duvall, T.L. and Harvey, J.W. 1983, *Nature* 302, 24
- Grec, G., Fossat, E. and Pomerantz, M., 1980, *Nature* 288, 541
- Grec, G., Fossat, E. and Pomerantz, M., 1983, *Solar Physics* 82, 55
- Rhodes, E.J., Jr., Ulrich, R.K. and Simon, G.W. 1977, *Ap. J.* 218, 901.
- Vandakurov, Y.U. 1967, *Astro. Zhurnal* 44, 768

SPECTRAL INFORMATION FROM GAPPED DATA:  
A COMPARISON OF TECHNIQUES

J. R. Kuhn  
Physics Department  
Princeton University

ABSTRACT

Three methods for estimating the power spectrum of gapped time series are briefly discussed. Numerical experiments, using several synthetic data sets, are used to compare these techniques.

1. INTRODUCTION

The FFT is widely used to estimate power spectra of continuous signals evenly sampled on discrete domains. This paper briefly discusses the problem of finding power spectra on unevenly sampled domains — in particular a regularly spaced domain with gaps. For example, the analysis of the ACRIM solar bolometric intensity data, obtained with a 3/5 on and 2/5 off duty cycle of approximately 100 minutes, would benefit from the techniques discussed below. We report on the comparative effectiveness of three different analysis techniques applied to synthetic data generated on gapped domains. Since the details of the analysis techniques have been presented in the literature only a general overview of the methods will be discussed here.

2. THE SPECTRAL DECOMPOSITION PROBLEM

Define  $t_j$ ,  $j = 0, \dots, N-1$  as a time domain with the  $t_j$  integers and  $t_0 = 0$ , and  $d_j$  is the corresponding data sequence again with  $j = 0, \dots, N-1$ . If we have reason to believe that the signal is well represented by a harmonic expansion of the form

$$d(t) = \sum_{j=0}^{M-1} a_j e^{i\omega_j t} \quad (1)$$

then our task is to determine the  $a_j$  from the data sequence  $d_j$ . Note that the  $N$ -tuple of data points  $(d_0, \dots, d_{N-1})$  is a vector  $\vec{d}$  in an  $N$ -dimensional complex vector space. If we write  $S_k = (1, e^{i\omega_k t_1}, \dots, e^{i\omega_k t_{N-1}})$  then Equ. (1) can be expressed in the form

$$\vec{d} = \sum_{k=0}^{M-1} a_k \vec{S}_k \quad (2)$$

Taking the usual scalar product of both sides of Equ. (2) with a given complex conjugate harmonic vector  $\vec{S}_j^*$  gives

$$D_j \equiv \vec{d} \cdot \vec{S}_j^* = \sum_{k=0}^{M-1} a_k \vec{S}_k \cdot \vec{S}_j^* \equiv \sum_{k=0}^{M-1} a_k C_{kj} \quad (3)$$

The matrix equation to be solved for the spectral coefficients  $a_j$  is then

$$D_j = C_{jk} a_k \quad (4)$$

where repeated indices are summed and  $D_j$  and  $C_{jk}$  are determined entirely by  $t_j$ ,  $d_j$  and  $\omega_j$ . This is the same equation one must solve to get the best multiparameter Least-Squares estimate for the  $a_j$ .

On an even domain  $t_j = j$  and with  $\omega_k = 2\pi k/N$  then  $C_{kj} = N\delta_{kj}$ , so that we get

$$a_k = \frac{1}{N} \sum_{j=0}^{N-1} \exp(-i2\pi kj/N) d_j$$

the familiar expression for a discrete Fourier Transform.

On an uneven domain in general  $t_j \neq j$  and we don't know how to choose the  $\omega_k$  so that  $C_{kj}$  is diagonal. If we let  $\omega_k = 2\pi k/N$  then  $C_{kj}$  is not diagonal. The uneven domain adversely affects the spectrum calculated by naively treating the gaps as zeros in two ways: First, noise in each coefficient  $a_k$  is larger than it would be if the data points were distributed without gaps, and second, the off diagonal elements in  $C_{jk}$  alias the signal peaks to other frequencies.

There are two approaches to solving this problem: 1) the data can be interpolated in the gaps and then the transform calculated on the regular domain, or 2) efficient numerical techniques may be applied to directly invert Equ. (4).

### 3. SOLUTIONS TO THE PROBLEM

An interpolation scheme has been widely used in the geophysics community that assumes the data at any time  $t_j$  can be expressed as a linear combination of previous data values plus a noise contribution (Fahlman and Ulrych, 1982 and references). The procedure I have used below is an adaptation of the Fahlman and Ulrych algorithm. Essentially a best least squares fit of the form

$$d_j = \sum_{k=1}^L \gamma_k d_{j-k} + e_j \quad (5)$$

is determined by minimizing the summed squared error, e.g.  $\sum e_j^2$ , between data predicted by coefficients  $\gamma_k$  (applied forward and backward) and actual data values. After determining the filter coefficients,  $\gamma_k$ , from data between gaps the data values in the gaps are determined from a similar least squares solution (the data points in the gaps are now considered unknowns instead of the coefficients,  $\gamma_k$ ). Once data values in the gaps have been obtained the procedure can be iterated using these values to recalculate the filter coefficients, and then a new set of predicted data in the gaps.

The second approach to the problem involves inverting the matrix equation (4). Since typically this will involve inverting a complex matrix that has hundreds of elements in each row, this is not easily

done directly. The simplest solution is to ignore the off diagonal elements of  $C_{jk}$  (which are smaller in magnitude than the diagonal  $C_{jj} = N$ ). This essentially yields the periodogram solution, or the same result one would get by replacing the gaps with zeros and performing an FFT.

Another approach to solving Equ. (4) is discussed by Kuhn (1982). For regularly sampled data with gaps  $C_{jk}$  is a circulant matrix (Davis, 1979) and can be readily inverted, or if it's singular a "generalized" inverse is easily obtained. In the case where  $C_{jk}$  is not singular the solution discussed above has a simple description. To construct the solution from the spectral coefficients we first find the largest  $M < N$  such that for each integer,  $i$ , between 0 and  $M-1$  there is at least one time domain point,  $t_j$ , equal to  $i$  modulo  $M$ . Having found the largest  $M$  satisfying this condition each of the data points  $d_j$  corresponding to time points that are separated by some multiple of  $M$  are averaged together to get a modified data string of  $M$  values with corresponding time points from 0 to  $M-1$ . An FFT is then applied to this "swapped" and averaged time series to get spectral coefficients at frequencies  $\omega_j = 2\pi j/M$ . An equivalent way of viewing the data point swapping follows by noting that if the data is well represented by an equation of the form

$$d_j = \sum_{k=0}^{M-1} \exp(i2\pi jk/M) a_k \quad (6)$$

then letting  $j \rightarrow j + nM$  doesn't change this equality -- with this representation of the signal the data is periodic with period  $M$ .

#### 4. NUMERICAL EXAMPLES

Synthetic data was generated by adding normally distributed random numbers to pure sinusoids sampled on various domains. Three different approaches were used to estimate the power spectra: 1) immediate FFT with zeros in the gaps, 2) fitting a predictive filter (called an AR model) and interpolating into the gap, 3) inverting Equ. (4) by swapping and averaging data points, and then applying an FFT on  $M < N$  data points. In each domain  $t_N \approx 1000$  and there are typically 8 cycles of approximately 50% duty cycle. Thus for the "swapped" solutions  $M \approx N/2$  and the spectral resolution is half what the AR scheme produces (which fills in the gaps so the resolution is determined by the total duration of observations).

Figure 1 shows how the interpolation method works on a pure sinusoid with no noise. The domain consists of 77 sample points and 47 gap points for 8 cycles. The frequency of the sinusoid is  $1/\pi$ . The first iteration of the AR model uses a filter length equal to  $1/4$  the gap length. On the second pass the filter length is doubled. From Fig. 1 note that the 2nd pass doesn't substantially improve on the interpolated data from the 1st pass. It is also not much affected by subsequent iterations or variations in filter length between  $1/4$  and  $1/2$  of the gap length.

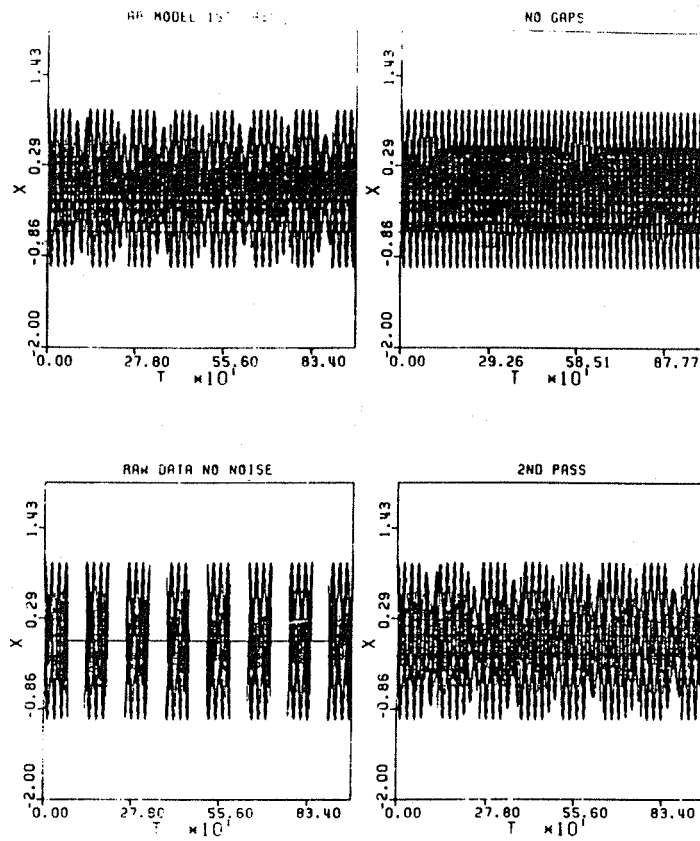


Figure 1: High Frequency AR Interpolation.

Figure 2 shows similar results for a sinusoid with period longer than the gap length. Since the filter is limited in length by the gaps it's not unexpected that it has a difficult time interpolating very low frequencies into the gap.

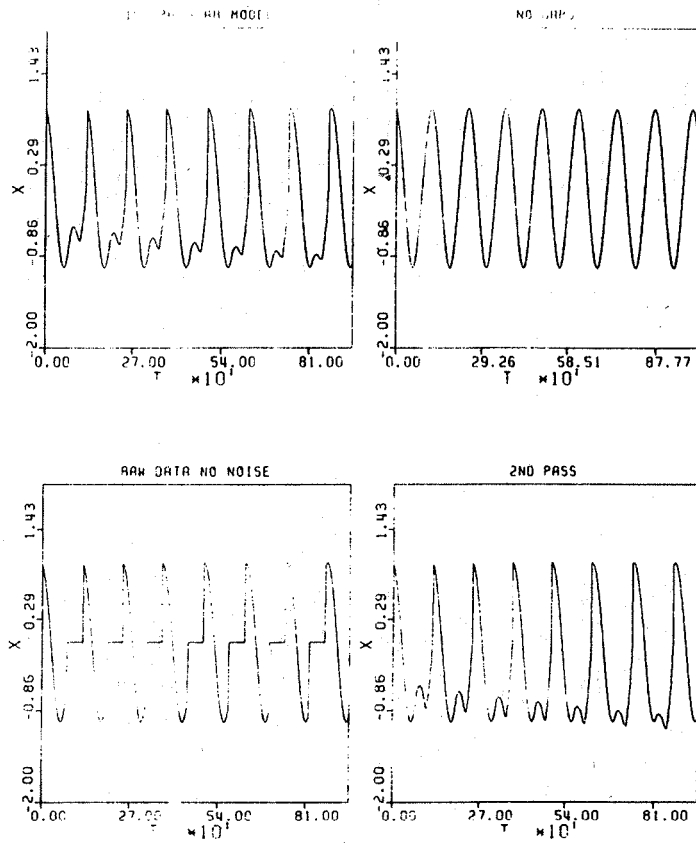


Figure 2: Low Frequency AR Interpolation.

Figure 3 shows the effect of additive noise on the AR scheme. The noise has  $\sigma = 2$  and the sinusoid frequency is  $3.2 \times 10^{-3}$ . Notice the noise in the gaps is not reproduced by the AR method and that the 2nd iteration shows a noticeable effect on the interpolated data.

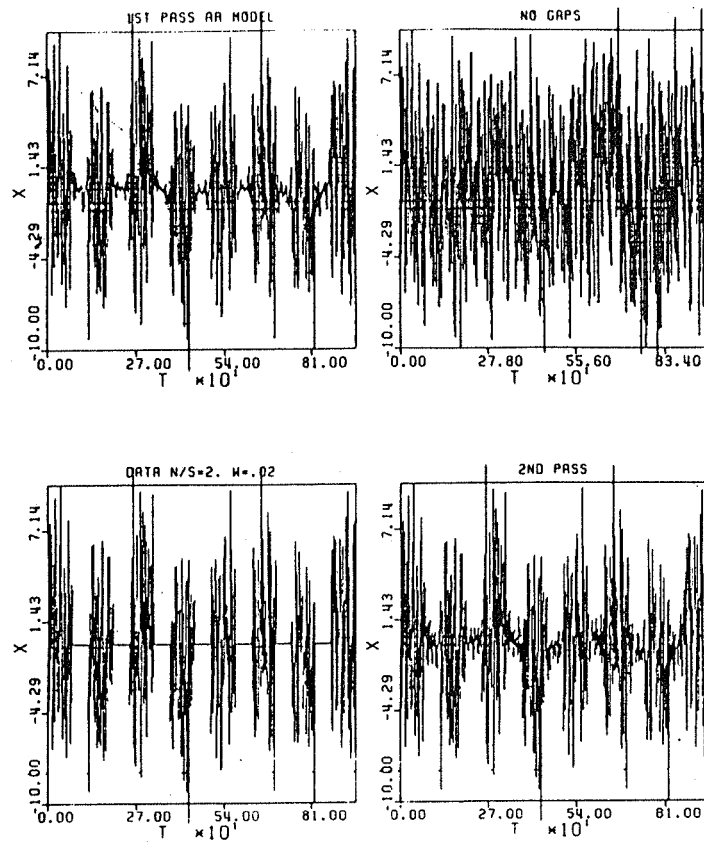


Figure 3: AR Interpolation on Noisy Data.



The next set of figures show how the power estimates are affected by the three schemes.

Figure 4 plots power estimates corresponding to the data from Fig. 1. The aliasing due to the gaps is apparent in the "Gap Power" segment of this Figure. Sidelobes are eliminated by both the AR and swap procedures. Note the amplitude estimate from the swap procedure is more accurate.

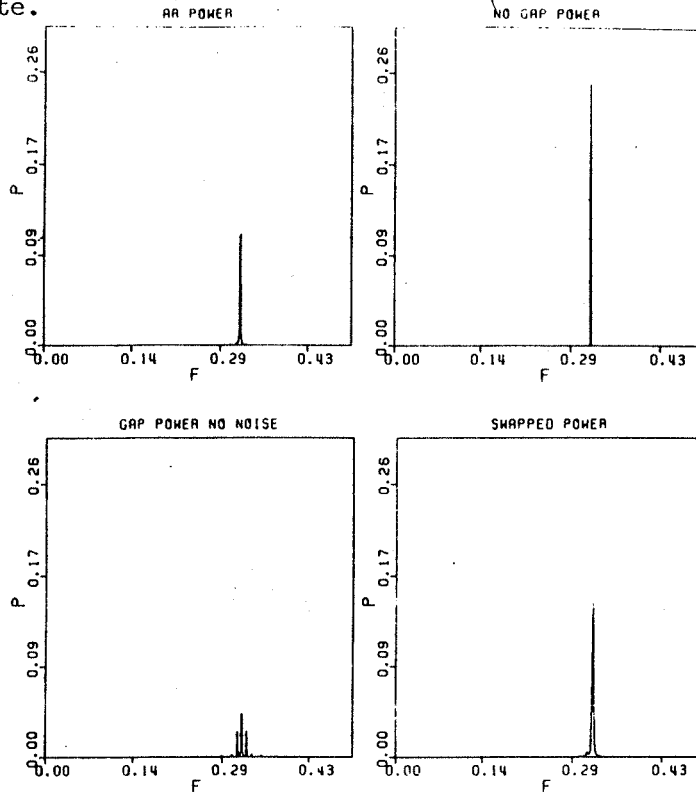


Figure 4: Power Estimates Using Fig. 2 Data.

Similar power estimates from the data in Fig. 2 are shown in Figure 5. Notice that the swapped spectrum is slightly worse than the AR spectrum, which is not itself an improvement over a straight FFT with zeros.

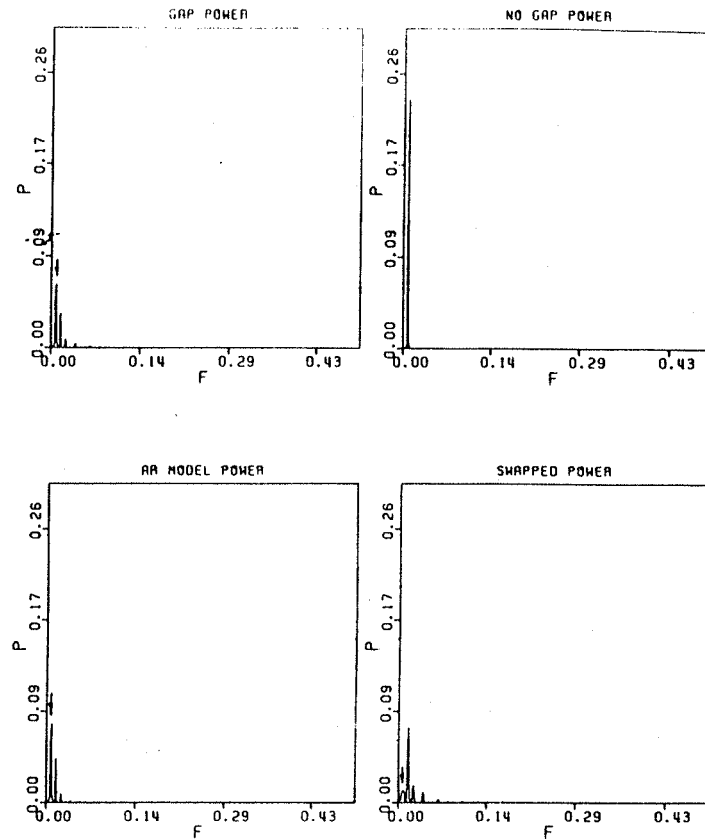


Figure 5: Power from Fig. 2 Data.

The next figure shows the effect of a more complicated spectrum --- in this case 4 unit amplitude sinusoids with random phases, and no additive noise. Figure 6 shows that the AR method does considerably worse than the swap scheme, which clearly resolves the four peaks from their aliases produced by the gapped domain.

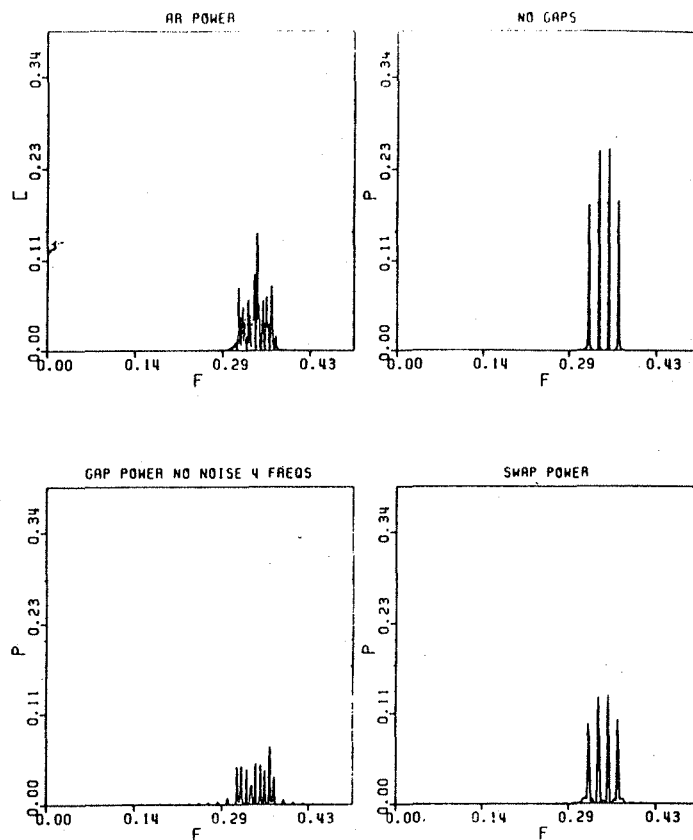


Figure 6: Power Estimates from Complex Signal.

The next two figures show how an increasing noise level affects the methods. In Figure 7 the noise is calculated with  $\sigma = 1$ . (The domain and signal are as in Fig. 1.) Both methods do much better than a transform with zeros. The data in Figure 8 were calculated with  $\sigma = 2$ . Here the signal peak is lost in both the AR and gapped FFT, but is determined by the swapped spectrum.

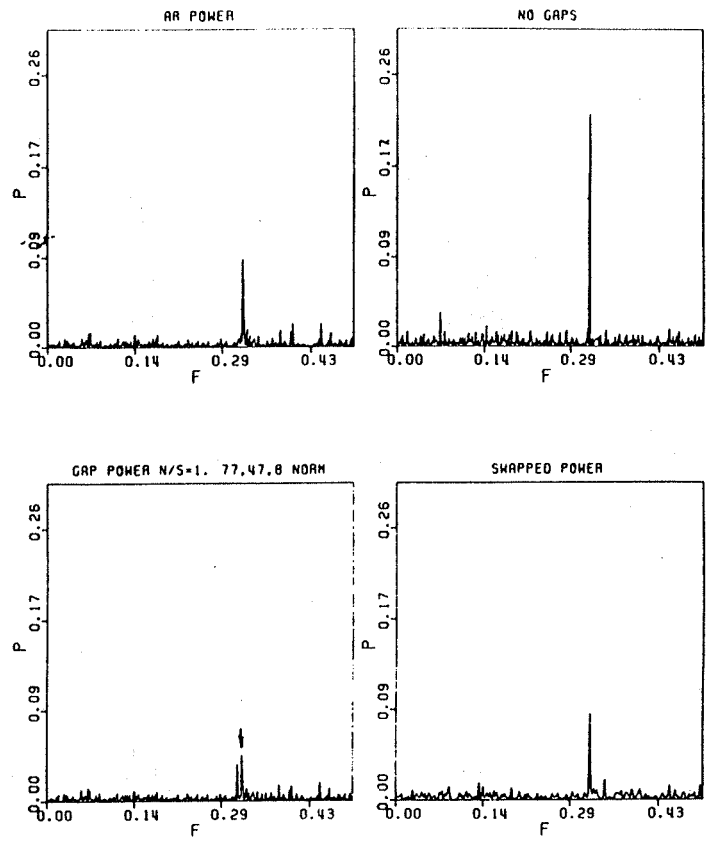


Figure 7: Power for Fig. 1 Data with Noise ( $N/S = 1$ ).

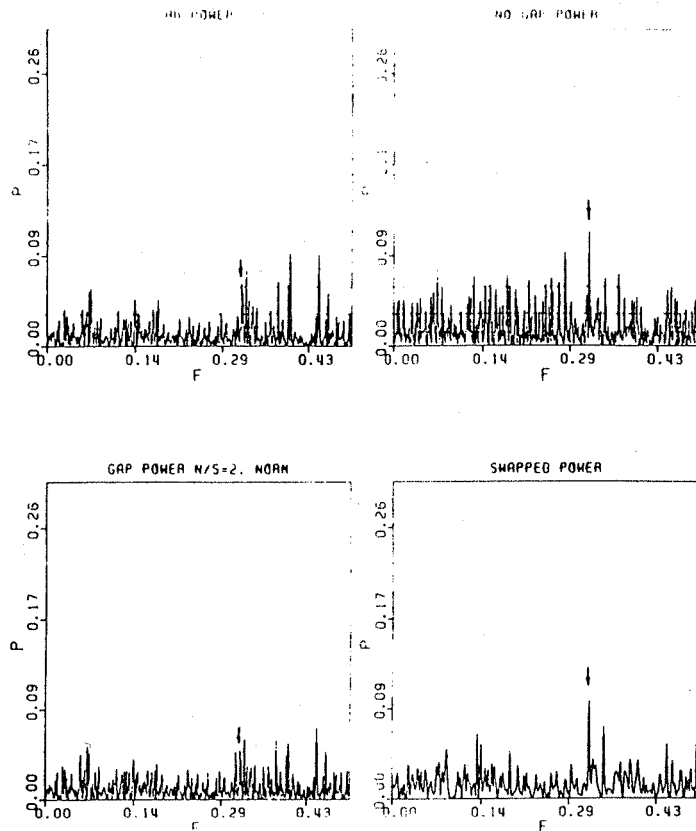


Figure 8: Power for Fig. 1 Data with Noise ( $N/S = 2$ ).

From these few numerical experiments it seems apparent that both the "swap" scheme and the AR method are useful for estimating power spectra on gapped domains of roughly 50% duty cycle and signal frequencies that are not too low. Furthermore it appears that the swap method works better than the AR approach for very noisy data or complicated spectra, but yields spectral resolution that may be more than a factor of 2 worse than an interpolation scheme. The full solution with a larger number of frequency points, discussed in the reference above, would improve the resolution.

#### REFERENCES

- Davis, P. J. 1979 Circulant Matrices (Wiley, New York).  
 Fahlman, G. G. and Ulrych, T. J. 1982, M.N.R.A.S. 199, 53.  
 Kuhn, J. R. 1982, Astron. J. 87, 196.



## METHODS FOR CLOUD COVER ESTIMATION

David L. Glackin, James R. Huning, Jeffrey H. Smith, Thomas L. Logan

Jet Propulsion Laboratory  
California Institute of Technology

### ABSTRACT

Several methods for cloud cover estimation are described relevant to assessing the performance of a ground-based network of solar observatories. These methods rely on ground and satellite data sources and provide meteorological or climatological information. Related studies completed and now in progress at the Jet Propulsion Laboratory are discussed.

### INTRODUCTION

One means of acquiring long-term observations of solar oscillations is the establishment of a ground-based network of solar observatories. Criteria for station site selection (listed in increasing order of difficulty) are: gross cloudiness, accurate transparency information, and seeing. As a first-order attempt at site selection, one or more candidate networks can be chosen on the basis of published climatological data. The average duty cycle for a network in terms of visibility/invisibility of the sun (a binary function) could then be computed. The discussion in this paper focuses on alternative methods for computing this duty cycle. The cycle, or alternatively a time history of solar visibility from the network, can then be input to a model, such as that developed by Hill at NCAR, to determine the effect of duty cycle on derived solar seismology parameters. Because personnel at JPL are involved in several studies of cloudiness from space, the authors have worked with the Ground Based Network Task Group (chaired by G. Newkirk, NCAR) to examine various means by which the duty cycle might be computed. Cloudiness, and to some extent transparency, can potentially be estimated from satellite data. Costs, however, are not inconsequential. To be meaningful, data from a five-year time span are desirable.

## GROUND-BASED INFORMATION

For network sites chosen to correspond to existing solar observatories, observing logs and flare patrols contain information on the visibility of the sun. Unfortunately, the observing logs from any solar observatory will be difficult to interpret in this respect, and a meaningful comparison of the logs from two or more observatories will be all but impossible. The quality of these records will vary according to the observer on duty, mechanical down time of the telescope and other effects. Recreating a timeline from flare patrol films could be attempted, as indeed it has for a 2 or 3 year period at Sacramento Peak Observatory (Gilliam, 1983). The possibility of creating a complete 5-year timeline from an entire network of observatories using this method is unlikely, due to the factors mentioned above. The major disadvantage is the restriction of the study only to existing solar observatories that have run flare patrols consistently for the same 5-year period. Such a constraint is unacceptable within the defined goals of this study.

Cloudiness data from standard meteorological stations at 3-hourly intervals are archived at Asheville, N.C., and by the World Meteorological Organization. Because of microclimatic variations the meteorology at the stations often bears little resemblance to the nearby mountaintop observatories. A good example is the Tucson airport site, which bears little resemblance to conditions on Kitt Peak. For some sites such as Big Bear, there is a nearby meteorology station. For other sites, an interpolation of surrounding meteorology stations may be appropriate. For still other sites, the use of satellite data may be called for. It should also be noted that availability of these data can be a function of the country of origin.

A feasible, relatively low-cost option is to generate synthetic timelines of cloudiness using climatological data, input to a Monte Carlo analysis model. Such models have a number of limitations, and in particular, the underlying assumptions must be examined with care. A climatological review of each site, including an analysis of persistence and other available data, could be used to provide the model input. The model could generate hourly values of solar visibility over a one-year period to yield a synthetic timeline of network performance. The timeline could illustrate periods when the network does or does not have line-of-sight to the sun due to clouds or darkness. This timeline would be used as input to the NCAR model. Such studies have been carried out at JPL in the past. This would be much less expensive than using satellite data, and in fact appears to be the least costly option. On the other hand, verification of the results must be done with care.



## SATELLITE DATA

Numerous studies have indicated that satellite platforms can be used to construct cloud cover data bases, although the majority of these studies involved only a small area and a short period of time. These studies include work by Hiser and Senn, 1980; Vonder Harr and Ellis, 1978; Hiser, 1978; Gautier, 1982; and Huning, et al, 1982. All of these studies involve the mapping of cloud cover or the evaluation of satellite platforms for the mapping of meso-scale insolation, with the primary parameter of interest being cloud cover (extent and persistence).

Because geosynchronous satellites and orbiting satellites provide continuous coverage over the earth's surface they are especially attractive for use in the development of a cloud cover data base. The primary complication in using satellite data for such a data base is the large volume of data that must be collected, analyzed and displayed. For example, for a nominal three year cloud cover data base of only the continental United States, with an average of 24 half-hourly acquisitions, some one billion bits must be registered, interpreted and converted to tabular files each day. This converts to approximately  $1 \times 10^{12}$  bits for a three year period. Clearly, the enormous amount of data that must be collected and evaluated would overwhelm most software-hardware configurations, especially if a system was designed for global coverage. To reduce the amount of data collected, analyzed and displayed the concept of a variable grid size can be used. In those areas of the world where cloud cover is homogeneous over large areas a large grid size can be used and the amount of data analyzed substantially reduced. In those areas where micro-climatic conditions dictate, a smaller grid size is required. In addition, in those areas of specific interest a small grid size can be incorporated.

There is no existing database of global meteorological cloud cover data, but the potential exists to construct it from various sources. Geosynchronous satellites which provide coverage of most of the globe are the U.S. GOES (Geostationary Operational Environmental Satellite), the ESA Meteosat and the Japanese GMS. The non-U.S. satellites have limited temporal coverage, and access to the data is problematical. Images from GOES East (Atlantic) and West (Pacific-see Figure 1) are obtained in the visible and near infrared with spatial resolution down to 0.9 by 0.9 km. These images are archived with half-hourly time resolution by the University of Wisconsin, under contract to NOAA. The images are archived on video tape, and routed through an analog-to-digital converter when written back onto magnetic tape for distribution. This often results in missing lines, missing blocks, garbled lines, and sometimes completely garbled images. Despite these problems, this satellite data base has proven to be most useful at JPL.

The NOAA 6, 7 and 8 polar orbiters are another potential source of cloud cover data. They acquire data twice daily with the AVHRR (Advanced Very High Resolution Radiometer) instrument. Obtaining navigation data to perform earth location of these images is difficult, and achieving registration with respect to earth coordinates to the pixel level, as would be required for sites such as Mt. Haleakala, is extremely difficult. Earth location to 10-15 km is probably feasible. The GOES images are easier to use and would result in greater accuracy although they do not cover the globe (GOES effectively misses latitudes  $>65$  degrees).

Cloudiness data are also available from the NOAA HIRS (High Resolution Infrared Sounder) instrument, discussed in the next section, and potentially from the DMSP (Defense Meteorological Satellite Program) Block 5 and the Tiros spacecraft.

Finally, Sadler and Kilonsky (1981) have produced relatively low-resolution maps of 7-year monthly averages of cloudiness, with latitudinal coverage from 40S to 50N based on the GOES data. This data base could prove more useful than the standard published climatological maps.

#### JPL STUDIES

At JPL, a pilot study of cloud cover for solar insolation assessment has recently been completed (Huning, et al, 1982). The study demonstrates the feasibility of deriving cloud cover statistics from satellite data. Four days of high-resolution GOES imagery over a limited geographical area were used. Elements of the study included decoding the GOES images supplied by the University of Wisconsin, geometrically registering the images, developing a normalization model to compensate for sun angle and creating a variable size georeference grid and a cloud/shadow model. Percent opacity was estimated on the basis of pixel intensity, and is accurate to about 20%. The potential exists to refine the accuracy to about 10%. However, this does not approach the transmissivity accuracy ultimately desired for solar seismology studies.

An ongoing study at JPL (M. Chahine) and Goddard (J. Susskind) involves the production of a 5-year world cloudiness atlas with a resolution of 2.5 by 3.0 degrees, based on the NOAA HIRS data. It will be presented as monthly averages. The task is to run for 3-4 years, and the data are to be archived in the JPL PODS (Pilot Ocean Data System). Motion pictures of cloudiness for selected regions are to be produced. Preliminary products from this study are shown in Figure 2, from the JPL Image Processing Laboratory (K. Hussey). The full atlas will not be available for some time, and the resolution will be relatively low for the purposes of the solar oscillations work.

Other related studies have been performed and are ongoing at JPL, including a pattern recognition method for cloud cover structure based on 2-D FFTs (Glackin, 1983). The authors may be contacted for further information.

#### CONCLUSION

There is evidence that mapping cloud cover from satellite platforms is within the range of requirements for siting ground-based solar observatories. As such requirements are refined, parallel developments could be (and are being) carried out in the cloud cover estimation area.

## REFERENCES

- Gauthier, C. (1982), "Mesoscale Insolation Variability Derived from Satellite Data", J. Appl. Meteor 21, 51.
- Gilliam, L (1983), private communication.
- Glackin, D. (1983), "Pattern Recognition for Convection Cell Structure Based on 2-D Fast Fourier Transforms", Bull. Am. Astron. Soc. 15, 720.
- Hiser, H. W. (1978), "Use of Satellites in Solar Applications", paper presented at International Symposium-Workshop on Solar Energy, 16-22 June 1978, Cairo, Egypt.
- Hiser, H. W. and Senn, H. V. (1980), "Mesoscale Mapping of Available Solar Energy at the Earth's Surface by Use of Satellites", Solar Energy 24, 129.
- Huning, J. R., Logan, T. L. and Smith, J. H. (1982), "Cloud Cover Estimation: Use of GOES Imagery in Development of Cloud Cover Data Base for Insolation Assessment", JPL Publication 82-102, November 30.
- Sadler, J. C. and Kilonsky, B. J. (1981), "Trade Wind Monitoring Using Satellite Observations", University of Hawaii, Department of Meteorology UHMET 81-01.
- Vonder Harr, J. H. and Ellis, J. S. (1978) "Determination of the Solar Energy Microclimate of the United States Using Satellite Data", Final Report NAS5-22372, July.

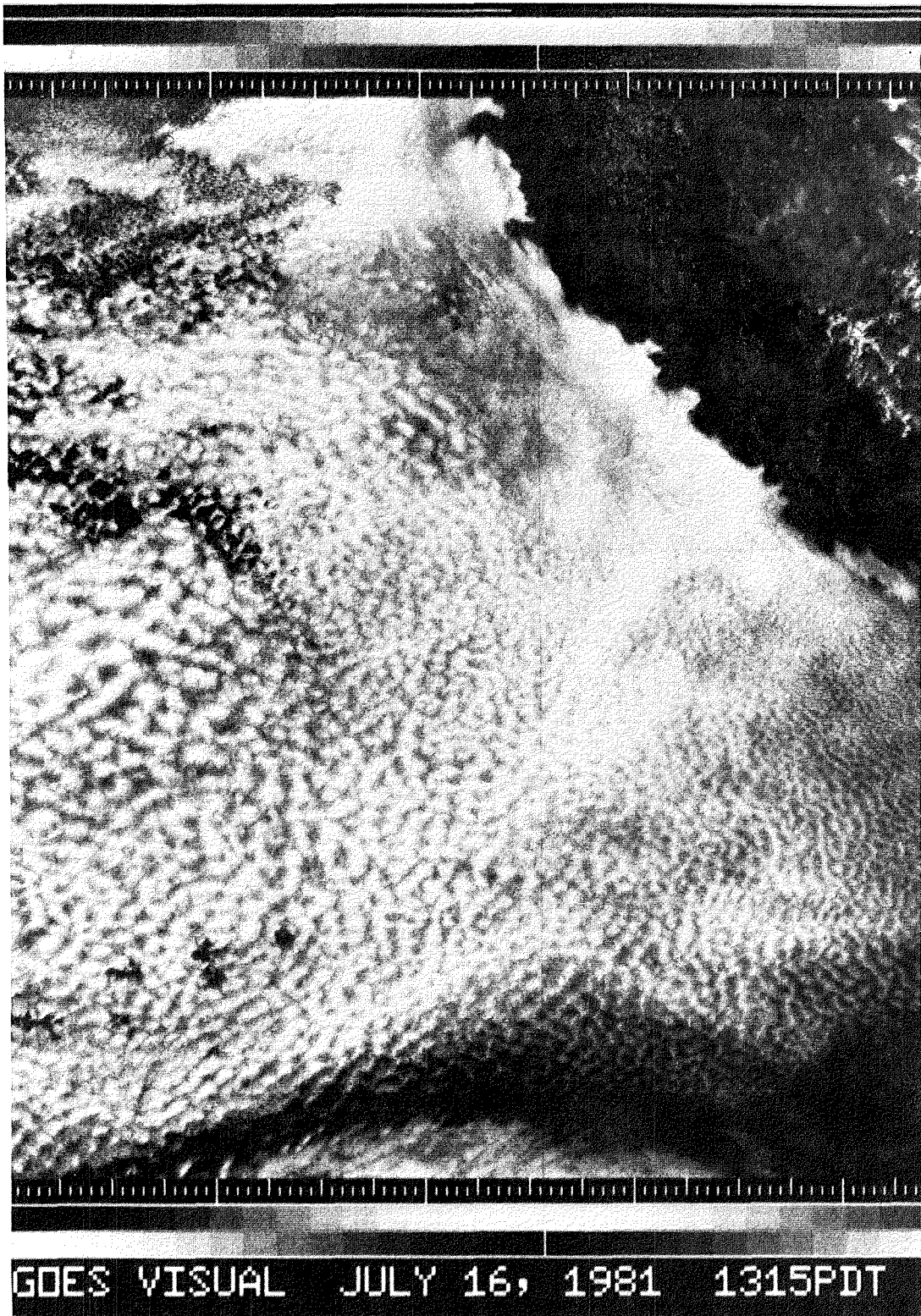
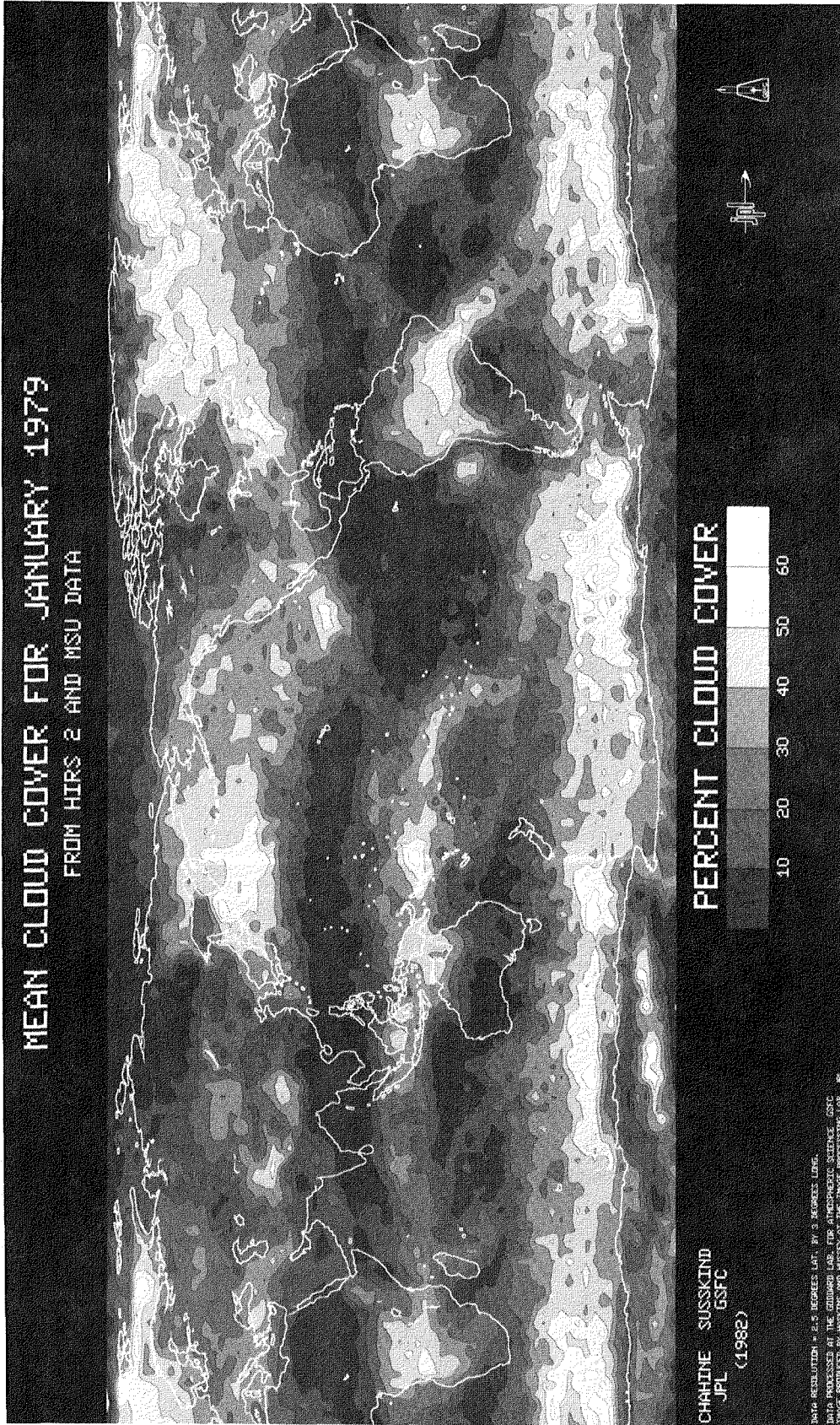


Figure 1. Two Mile Resolution GOES Visual Imagery off California Coast

Figure 2.



PART 4

EXCITATION, DAMPING AND ROTATIONAL SPLITTING OF THE OSCILLATIONS





## NUMERICAL SIMULATIONS OF CONVECTIVELY EXCITED GRAVITY WAVES

Gary A. Glatzmaier  
Theoretical Division  
Los Alamos National Laboratory  
Los Alamos, New Mexico

Magneto-convection and gravity waves are numerically simulated with a nonlinear, three-dimensional, time-dependent model of a stratified, rotating, spherical fluid shell heated from below. A Solar-like reference state is specified while global velocity, magnetic field, and thermodynamic perturbations are computed from the anelastic magnetohydrodynamic equations. Convective overshooting from the upper (superadiabatic) part of the shell excites gravity waves in the lower (subadiabatic) part. Due to differential rotation and Coriolis forces, convective cell patterns propagate eastward with a latitudinally dependent phase velocity. The structure of the excited wave motions in the stable region is more time-dependent than that of the convective motions above. The magnetic field tends to be concentrated over giant-cell downdrafts in the convective zone but is affected very little by the wave motion in the stable region.

### 1. INTRODUCTION

I would like to illustrate with numerical simulations how complex the global velocity and magnetic fields must be in the Sun, and how their structure and evolution in the stable region differ from that in the unstable, convective region. The dynamic dynamo model and the numerical method are described in Glatzmaier (1983). The anelastic MHD equations reduce to a 17th-order system of equations with each of the six dependent variables expanded in 1024 spherical harmonics and 17 Chebyshev polynomials. A semi-implicit time-integration scheme is employed. The anelastic approximation filters out acoustic waves but not gravity or convective modes.

A typical snapshot of the simulated motions is shown in Figure 1 where the mass flux (velocity times density) is plotted in the equatorial plane. The top and bottom boundaries correspond to 93% and 56% of the Solar radius, respectively. There are seven pressure scale-heights across the shell. The  $(\nabla - \nabla_{AD})$  profile in Figure 2 defines the superadiabatic and subadiabatic regions. (The stable region in a

standard Solar mixinglength model is much more subadiabatic than it is in this model. See further comments below.) Notice in Figure 1 how convective motions overshoot a short way into the stable (subadiabatic) region. This overshooting excites small amplitude gravity waves in the differentially rotating fluid.

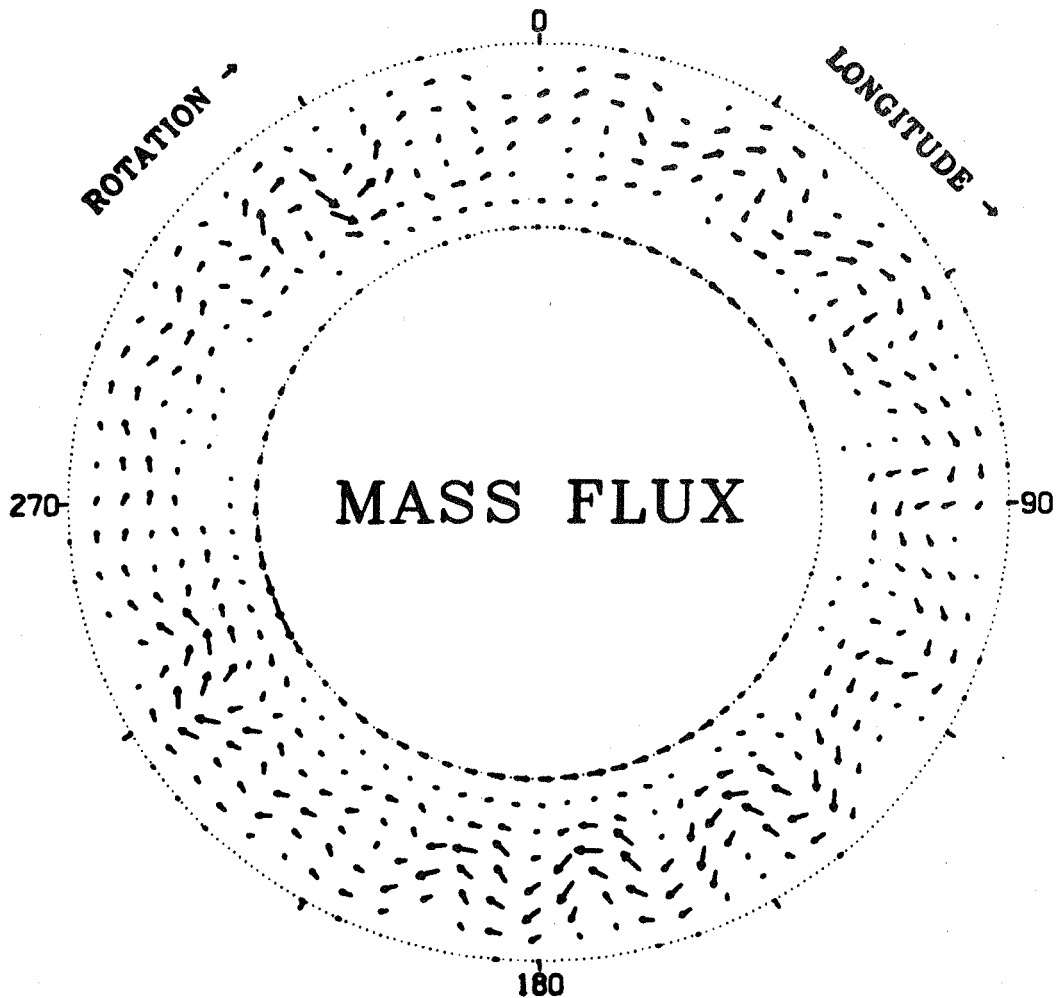


Figure 1. Mass flux vectors plotted in the equatorial plane viewed toward the north.

Generally, buoyancy forces do positive work in the superadiabatic region by driving convection; while, in the subadiabatic region, they do negative work by damping gravity waves. This is illustrated in Figure 2 where the horizontally averaged work done by buoyancy is plotted vs. radius. The negative buoyancy work in the stable region is small because the perturbations are small relative to those in the convection zone. Notice how buoyancy does positive work a short way into the subadiabatic region. This is due to the overshooting of sinking fluid that remains heavier than the surroundings for a short distance into the subadiabatic region. The negative work done by buoyancy near the top of the convection zone is required to help decelerate rising fluid and accelerate sinking fluid (Glatzmaier and Gilman 1981b).

We will examine the structure and evolution of the velocity and magnetic fields - first in the unstable (superadiabatic) region and then in the stable (subadiabatic) region.

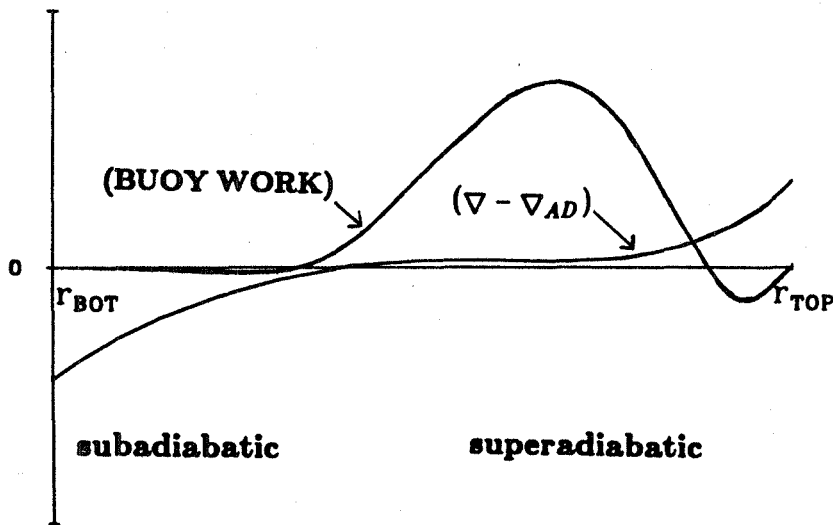


Figure 2. Horizontally averaged  $(\nabla - \nabla_{AD})$  and buoyancy work density per time plotted vs. radius.

## 2. THE UNSTABLE REGION

In Figure 3, the radial component of the velocity is plotted (relative to the rotating frame of reference) in a spherical surface just below the top boundary at four different times. The snapshots were taken a week apart. As suggested by linear anelastic calculations (Glatzmaier and Gilman 1981b), north-south rolls are preferred because, for these, Coriolis forces are more easily balanced by pressure gradient forces. However, nonlinear, multimode calculations are required to study the complicated structure and evolution of the convective cells.

One can see, by examining these snapshots, how the cell pattern in the equatorial region propagates eastward relative to the patterns in the polar region. As a result, an updraft region is periodically sheared apart at mid-latitude and subsequently joined to the next updraft region. The eastward cell velocity in the equatorial region is about 10% of the average Solar rotation rate.

There are two reasons for this type of evolution. The obvious one is the nonlinear effect of the latitudinal differential rotation. That is, the north-south rolls become deformed because the equatorial region rotates faster due to the transport of angular momentum (Gilman 1977, Glatzmaier and Gilman 1982). The other reason is a linear effect due to Coriolis forces (Glatzmaier and Gilman 1981a). As rising fluid in a north-south roll expands, Coriolis forces cause it to rotate in the opposite sense of the global rotation generating negative local vorticity. Positive local vorticity is generated when sinking fluid contracts. Consequently, since positive vorticity exists in north-south rolls that are east of updrafts and negative vorticity in rolls east of downdrafts, the phase of the north-south rolls propagates eastward. Since this effect is greatest where gravity is perpendicular to the rotation axis, the resulting eastward phase velocity is maximum in the equatorial region.

Now we examine, in Figure 4, the corresponding structure and evolution of the radial component of the magnetic field in the same spherical surface at the same four times. In this surface the magnetic energy density is approximately one thousand times smaller than the kinetic energy density. As a result, both magnetic field polarities tend to be concentrated over the downdrafts of the giant cells due to the convergence of horizontal flow (Glatzmaier 1983). This can be seen by carefully comparing Figures 3 and 4. Since the peak radial components of the magnetic field exist at mid-latitude, they do not experience the large eastward phase velocity of the convective cells in the equatorial region. However, by close examination, one can see how the magnetic field structure changes slightly as the convective rolls are sheared at mid-latitude.

### 3. THE STABLE REGION

The structure and evolution of the velocity and magnetic fields in the stable region are quite different than what has just been described for the unstable region. The radial component of the velocity is plotted in Figure 5 for the same four times but in a spherical surface just above the bottom boundary. Since here buoyancy is a restoring force, the velocity consists of oscillatory wave motions. However, these are not simple linear gravity modes but highly structured, time-dependent, nonlinear waves which are continually being excited by convective overshooting and affected by Coriolis and Lorentz forces on a differentially rotating fluid background in spherical geometry.

The corresponding radial component of the magnetic field, plotted in Figure 6, shows little change with time and very little correlation with the velocity field. This is probably because the oscillatory fluid motions deep in the stable region do not have horizontal convergence properties as do the convective motions in the unstable region. Also, the magnetic energy density at this depth is only ten times smaller than the kinetic energy density.

### 4. SUMMARY

These numerical simulations were presented to illustrate how complicated the structure and evolution of the velocity and magnetic fields must be in the Sun. They illustrate the latitudinally dependent eastward propagation and resulting shearing of the north-south rolls. Certainly this makes the observation of Solar giant cells difficult, especially when the data is averaged over several weeks in order to reduce the small-scale Solar noise (Howard and LaBonte 1980, Gilman and Glatzmaier 1980). On the other hand, the simulated large-scale magnetic fields, concentrated over giant-cell downdrafts near the surface at mid-latitude, change very little with time and resemble large-scale Solar magnetic field observations (Howard 1977).

The simulations also illustrate how complicated gravity wave motions are in the stable region and how they differ from convective motions in the unstable region. However, the simulations were not meant to predict periods of the excited gravity modes. The periods, which are of the order of a few weeks, probably are not realistic for two reasons. The model's impermeable bottom boundary forces an artificial node and enhances the higher order spherical harmonic modes (smaller scales). In addition,  $(\nabla \cdot \nabla_{\perp})$  in the stable region is much smaller than it is in a standard Solar model; consequently, the gravity wave periods are much larger than they would be if the stable region were more subadiabatic.

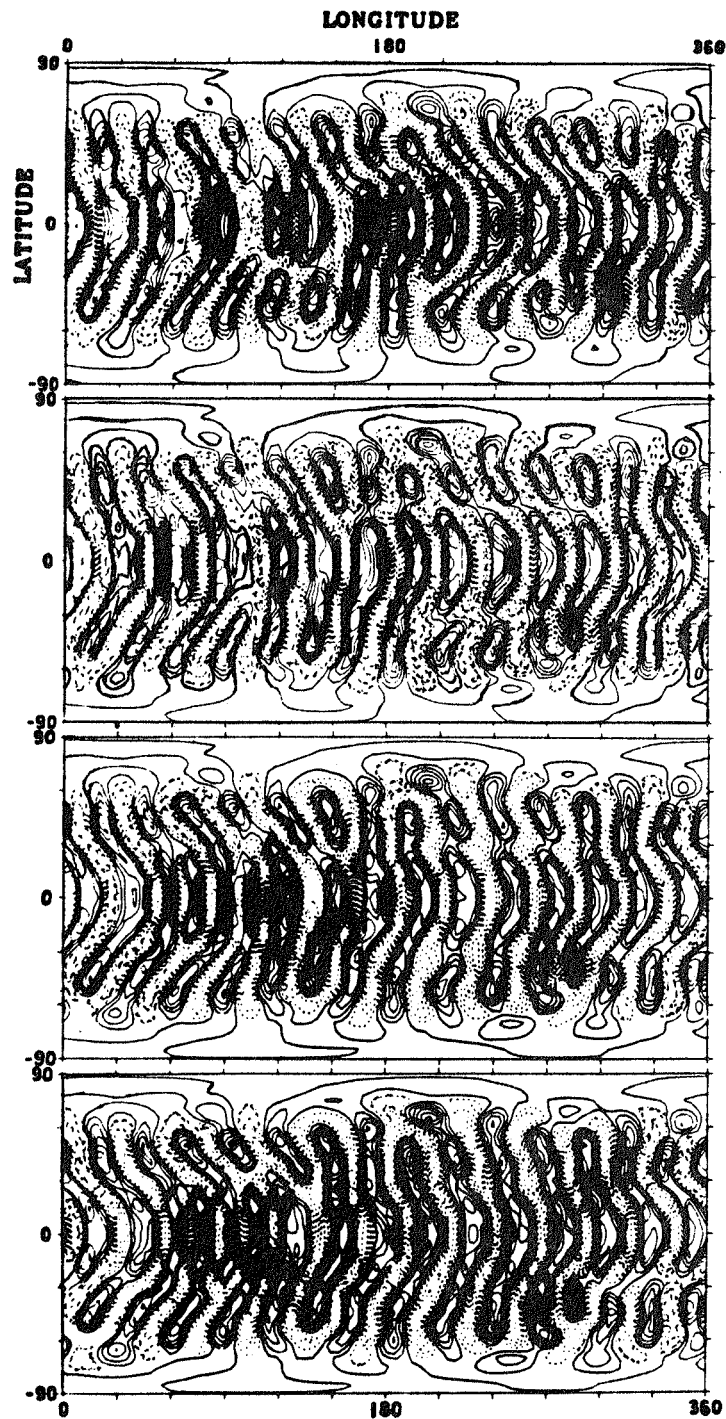


Figure 3. Radial components of velocity plotted in a spherical surface just below the top boundary at one week intervals. Solid (broken) contours represent upward (downward) velocity.

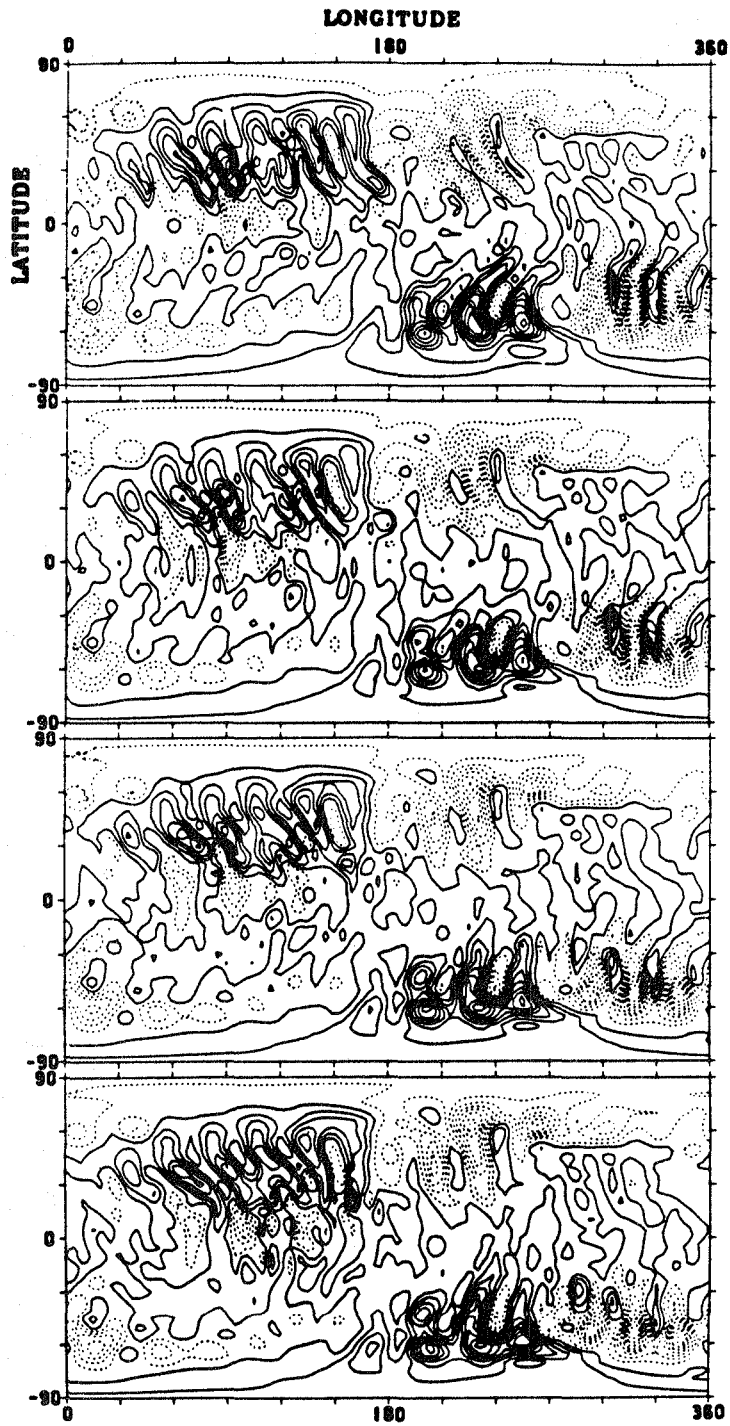


Figure 4. As in Figure 3, but for the radial component of the magnetic field.

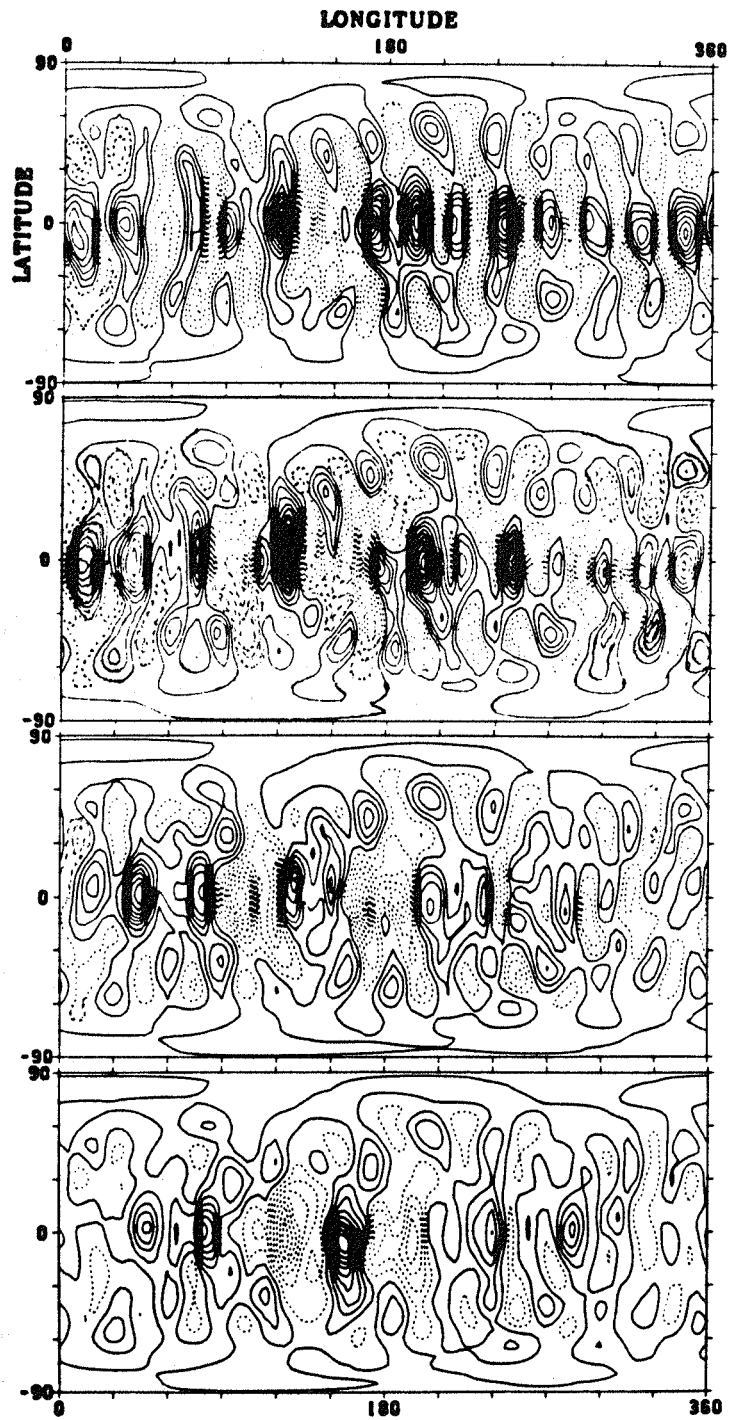


Figure 5. As in Figure 3, but for a spherical surface just above the bottom boundary.



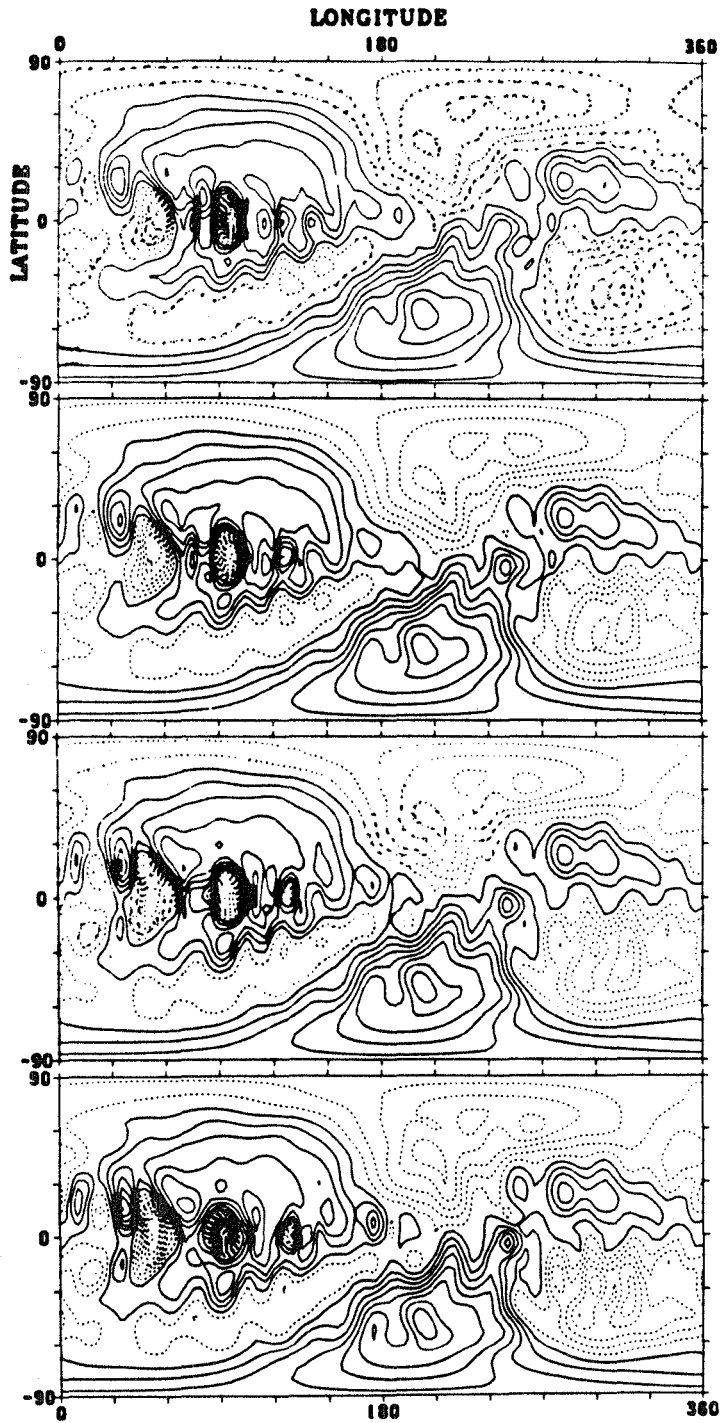


Figure 6. As in Figure 5, but for the radial component of the magnetic field.

A new version of the code is being developed that will model an entire sphere, employ time-dependent sub-grid scale eddy diffusivities, and incorporate a  $(\nabla - \nabla_{AD})$  profile based on a standard Solar model. This new code should be able to predict more realistic gravity wave amplitudes and periods. However, the present results illustrate how difficult it would be to observe these highly structured, time-dependent gravity modes, especially if the interpretation is based on linear models that do not account for convective motions, Coriolis forces, Lorentz forces, or differential rotation.

#### REFERENCES

- Gilman, P. A.: 1977, Geophys. Astrophys. Fluid Dyn., 8, 93.  
Gilman, P. A. and Glatzmaier, G. A.: 1980, Ap.J., 241, 793.  
Glatzmaier, G. A.: 1983, J. Comp. Phys., in press.  
Glatzmaier, G. A. and Gilman, P. A.: 1981a, Ap.J. Suppl., 45, 381.  
Glatzmaier, G. A. and Gilman, P. A.: 1981b, Ap.J. Suppl., 47, 103.  
Glatzmaier, G. A. and Gilman, P. A.: 1982, Ap.J., 256, 316.  
Howard, R.: 1977, Ann. Rev. Astron. Astrophys., 15, 153.  
Howard, R. and LaBonte, B. J.: 1980, Ap.J., 239, 738.

## ON THE INFLUENCE OF TURBULENT MOTIONS ON NON-RADIAL OSCILLATIONS

B. R. Durney  
Sacramento Peak Observatory\*

The effect of turbulent motions on oscillations is studied, considering only the coupling between turbulent and oscillatory velocities. In this case, the turbulence affects the oscillations through the Reynolds stresses in the momentum equation for the pulsations. A simple model of turbulence is adopted to evaluate these Reynolds stresses and the perturbed eigenfrequencies are expressed as a function of certain averages of the turbulent velocities.

### INTRODUCTION

The study of time-dependent convection (Unno 1967; Gabriel et al. 1975; Gough 1977; Deupree 1977; Xiong 1977; Keeley 1977; Goldreich and Keeley 1977; Baker and Gough 1979; Saio 1980; Gonczi and Osaki 1980) is an important subject for the understanding of pulsating stars with convection zones.

Several of the above papers study the generalization of the mixing length theory of convection to pulsating stars and the influence of pulsation on the velocity and temperature fluctuations must be considered.

The aim of this paper is far more modest; we retain the influence of the oscillations on the turbulent velocities only, and neglect it on the turbulent values of the pressure, density and temperature. Therefore, the effect of the turbulent motions on the oscillations manifests itself solely through the mean Reynolds stresses in the momentum equation for the pulsations. The definition of mean quantity is based on time averages: it is assumed that the oscillations remain coherent for a time much longer than the life-time of the turbulent eddies. To calculate the mean Reynold stresses we adopt a very simple model of

---

\* Operated by the Association of Universities for Research in Astronomy, Inc. under contract AST 78-17292 with the National Science Foundation

turbulence; namely, we assume that the turbulent velocities remain unchanged for a turnover time and then suddenly change to an uncorrelated velocity field. The perturbed equation for the oscillations is solved by a variational principle and the perturbed eigenfrequencies,  $\omega'$ , are expressed as a function of  $\omega_0$ , the turbulent eddies' lifetime and certain averages of the turbulent velocities.

## THEORY

### a) The Momentum Equation

The equation for conservation of momentum can be written (we neglect rotation):

$$D\tilde{u}/Dt = \nabla\phi - \rho^{-1}\nabla p, \quad (1)$$

where  $\tilde{u}$ ,  $\phi$ ,  $\rho$ , and  $p$  are, respectively, the fluid velocity, the gravitational potential, the density and the pressure. The convective rate of change operator,  $D/Dt$ , is defined by

$$D/Dt = \partial/\partial t + \tilde{u} \cdot \nabla. \quad (2)$$

For  $\tilde{u}$ ,  $p$ ,  $\rho$ , and  $\phi$  we can write

$$\tilde{u} = \bar{u} e^{i\omega t} + \tilde{u}', \quad (3a)$$

$$p = \bar{p} e^{i\omega t} + p_0 + p'_0, \quad (3b)$$

$$\rho = \bar{\rho} e^{i\omega t} + \rho_0 + \rho'_0, \quad (3c)$$

$$\phi = \phi_0. \quad (3d)$$

In Eqs. (3), the quantities with a bar are oscillating quantities;  $p_0$ ,  $\rho_0$  and  $\phi_0$  are the pressure, density and gravitational potential in the absence of turbulence and oscillations;  $\tilde{u}'$ ,  $p'_0$ ,  $\rho'_0$  are the turbulent velocity, pressure and density. It is clear from Eq. (3d) that we neglect the effect of the oscillation and turbulent motions on the gravitational potential. The subscript '0' in  $p'_0$  and  $\rho'_0$  indicates that we also neglect the effect of the oscillation on the turbulent values of the pressure and density, i.e., we assume, for example, that  $p'(\bar{u} \neq 0) = p'(\bar{u} = 0) = p'_0$ . We stress, on the other hand, that we do not neglect the effect of the oscillation on  $\tilde{u}'$ . In fact the derivation of the equation relating  $\tilde{u}'$  and  $\tilde{u}$  is one of the main objects of this paper (the subscript '0' designates, here and elsewhere, quantities in the absence of oscillations).

We expand  $1/\rho$  as follows (see Eq. (3c)):

$$\frac{1}{\rho} = \frac{1}{\rho_0} \left( 1 - \frac{\bar{\rho} e^{i\omega t}}{\rho_0} - \frac{\rho'_0}{\rho_0} \right). \quad (4)$$

With the help of the unperturbed equation (Eq. (5b) below) and of Eqs. (3) and (4) it is straightforward to show that Eq. (1) can be written as follows:

$$\begin{aligned}
 & i\omega \bar{u} e^{i\omega t} + \partial \underline{u}' / \partial t + e^{i\omega t} \bar{u} \cdot \nabla \underline{u}' + e^{i\omega t} \underline{u}' \cdot \nabla \bar{u} \\
 & + \underline{u}' \cdot \nabla \underline{u}' - \langle \underline{u}'_0 \cdot \nabla \underline{u}'_0 \rangle_E = - e^{i\omega t} \rho_0^{-1} \nabla \bar{p} - \rho_0^{-1} \nabla p'_0 \\
 & + e^{i\omega t} \frac{\bar{\rho}}{\rho \rho_0} \nabla p_0 + e^{i\omega t} \frac{\bar{\rho}}{\rho} \rho_0^{-2} \nabla p'_0 \\
 & + \rho_0' \rho_0^{-2} \nabla p_0 + e^{i\omega t} \rho_0' \rho_0^{-2} \nabla \bar{p} + \rho_0' \rho_0^{-2} \nabla p'_0 - \rho_0^{-2} \langle \rho_0' \nabla p'_0 \rangle_E \quad (5a)
 \end{aligned}$$

$$- \langle \underline{u}'_0 \cdot \nabla \underline{u}'_0 \rangle_E - \frac{1}{\rho_0} \nabla p_0 + \rho_0^{-2} \langle \rho_0' \nabla p'_0 \rangle_E + \nabla \phi_0 = 0. \quad (5b)$$

In Eqs. (5) the bracket denotes an ensemble average and in Eq. (5a) we have neglected terms quadratic in the oscillations. We assume now that the oscillations remain coherent for a time much longer than the lifetime of the turbulent eddies. We multiply Eq. (5a) by  $e^{-i\omega t}/T$  and integrate in time from zero to  $T$ . In the limit  $T \rightarrow \infty$  we obtain

$$i\omega \bar{u} + e^{-i\omega t} \langle \underline{u}' \cdot \nabla \underline{u}' \rangle = - \frac{1}{\rho_0} \nabla \bar{p} + \frac{\bar{\rho}}{\rho_0^2} \nabla p_0, \quad (6)$$

with

$$\langle \underline{u}' \cdot \nabla \underline{u}' \rangle = e^{i\omega t} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T \underline{u}' \cdot \nabla \underline{u}' e^{-i\omega t} dt. \quad (7)$$

The justification of Eq. (6) is straightforward, since it is clear that an integral of the form  $\frac{1}{T} \int_0^T L' dt$  where  $L'$  is linear in the turbulent quantities, vanishes for large  $T$ 's. It is also clear that the same holds true for  $\frac{1}{T} \int_0^T \frac{\rho_0' \nabla p'_0}{\rho_0^2} e^{-i\omega t} dt$ . The quantity  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$ , on the other hand, differs from zero; this is due to the influence of the oscillation on the turbulent motions (see Eq. (9) below).

Equation (6) plays the role of the mean equation and we proceed now to evaluate  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$ . We multiply Eq. (6) by  $e^{i\omega t}$  and subtract it from Eq. (5a). It is readily found that the fluctuating momentum equation is given by

$$\begin{aligned}
 & \partial \underline{u}' / \partial t + e^{i\omega t} \bar{u} \nabla \underline{u}' + e^{i\omega t} \underline{u}' \cdot \nabla \bar{u} \\
 & + \underline{u}' \cdot \nabla \underline{u}' - \langle \underline{u}'_0 \cdot \nabla \underline{u}'_0 \rangle_E - \langle \underline{u}' \cdot \nabla \underline{u}' \rangle
 \end{aligned}$$

$$\begin{aligned}
&= -\rho_0^{-1} \nabla p'_0 + e^{i\omega t} \frac{\bar{\rho}^{-2}}{\rho \rho_0} \nabla p'_0 + \rho_0' \rho_0^{-2} \nabla p_0 \\
&\quad + e^{i\omega t} \rho_0' \rho_0^{-2} \bar{\rho} + \rho_0' \rho_0^{-2} \nabla p'_0 - \rho_0^{-2} \langle \rho_0' \nabla p'_0 \rangle_E. \tag{8}
\end{aligned}$$

Let  $\underline{u}'_0$  be the turbulent velocities in the absence of oscillations. The equation for  $\underline{u}'_0$  is obtained by setting  $\bar{u}, \bar{\rho}$ , and  $\bar{p}$  equal to zero in Eq. (8). We subtract the equation for  $\underline{u}'_0$  from Eq. (8), and replace  $\underline{u}'$  by  $\underline{u}'_0$  in the terms containing the perturbation  $\bar{u}$ . We obtain

$$\begin{aligned}
\frac{\partial \underline{u}'}{\partial t} &= \frac{\partial \underline{u}'_0}{\partial t} - e^{i\omega t} \bar{u} \cdot \nabla \underline{u}'_0 - e^{i\omega t} \underline{u}'_0 \cdot \nabla \bar{u} - \underline{u}' \cdot \nabla \underline{u}' + \underline{u}'_0 \cdot \nabla \underline{u}' + \langle \underline{u}' \cdot \nabla \underline{u}' \rangle \\
&+ e^{i\omega t} \frac{\bar{\rho}^{-2}}{\rho \rho_0} \nabla p'_0 + e^{i\omega t} \rho_0' \rho_0^{-2} \bar{\rho}. \tag{9a}
\end{aligned}$$

Neglecting quadratic terms in the turbulent quantities, Eq. (9a) reduces to

$$\begin{aligned}
\frac{\partial \underline{u}'}{\partial t} &= \frac{\partial \underline{u}'_0}{\partial t} - e^{i\omega t} \bar{u} \cdot \nabla \underline{u}'_0 - e^{i\omega t} \underline{u}'_0 \cdot \nabla \bar{u} + e^{i\omega t} \frac{\bar{\rho}^{-2}}{\rho \rho_0} \nabla p'_0 \\
&\quad + e^{i\omega t} \rho_0' \rho_0^{-2} \bar{\rho}. \tag{9b}
\end{aligned}$$

It is clear that the term  $\underline{u}' \cdot \nabla \underline{u}' - \underline{u}'_0 \cdot \nabla \underline{u}'_0 - \langle \underline{u}' \cdot \nabla \underline{u}' \rangle$  in Eq. (9a) need not be small for fully developed turbulence. This approximation finds justification more readily in the case of random waves (see Moffat 1978, p. 156). It should be stressed, however, that this "first order smoothing" or "quasi-linear" approximation has been used successfully by Krause (1968) in the derivation of the  $\alpha$ -term in dynamo theory and lies at the basis of our understanding of the solar cycle. Equation (9b) relates the turbulent velocities in the presence and absence of oscillations and it is, therefore, the central equation of this paper, since it allows us to evaluate  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$  from Eq. (7). To evaluate this term we consider the following very simple model of turbulence: we assume that the turbulent velocities remain relatively unchanged for a turnover time,  $\tau$ , and then suddenly change to an uncorrelated velocity field. To illustrate the details of the derivation of  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$  it suffices to retain only the first two terms in the right-hand side of Eq. (9b). Therefore,

$$\underline{u}'(t) = \underline{u}'_0(t) - \int_0^t e^{i\omega t'} \bar{u} \cdot \nabla \underline{u}'_0 dt' \tag{10}$$

The value of the integration constant does not play a role in our calculations, and we have assumed in Eq. (10) that it vanishes. It is clear that  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$  contains the term  $\langle \underline{u}'_0 \cdot \nabla \underline{u}'_0 \rangle$  and two terms linear in  $\bar{u}$  (we neglect quadratic terms in  $\bar{u}$ ). Now, by definition,

$\langle \underline{u}'_0 \cdot \nabla \underline{u}'_0 \rangle = e^{i\omega t} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-i\omega t} \underline{u}'_0 \cdot \nabla \underline{u}'_0 dt$ . This quantity vanishes, since the frequency  $\omega$  is of no particular significance for the unperturbed turbulence. The two terms linear in  $\bar{u}$  are
 
$$- \langle \int_0^t e^{i\omega t'} (\bar{u} \cdot \nabla \underline{u}'_0(t')) dt' \cdot \nabla \underline{u}'_0(t) \rangle$$

$$- \langle (\underline{u}'_0(t) \cdot \nabla) \int_0^t e^{i\omega t'} \bar{u} \cdot \nabla \underline{u}'_0(t') dt' \rangle.$$

We proceed now to derive the first term designated hereafter by  $\underline{A}$ . By definition

$$\underline{A} = -\lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-i\omega t} \underline{f}(t) \cdot \nabla \underline{u}'_0(t) dt \quad (11a)$$

where

$$\underline{f}(t) = \int_0^t e^{i\omega t'} \bar{u} \cdot \nabla \underline{u}'_0(t') dt'. \quad (11b)$$

Our model of turbulence allows us to write

$$\underline{f}(t) = \sum_{n=0}^{N-1} \int_{n\tau}^{(n+1)\tau} e^{i\omega t'} \bar{u} \cdot \nabla \underline{u}'_0(t') dt'$$

$$+ \int_{N\tau}^t e^{i\omega t'} \bar{u} \cdot \nabla \underline{u}'_0(t') dt' = \frac{(e^{i\omega\tau} - 1)}{i\omega} \sum_{n=0}^{N-1} e^{i\omega n\tau} \bar{u} \cdot \nabla \underline{u}'_0(n)$$

$$+ \frac{e^{i\omega t} - e^{i\omega N\tau}}{i\omega} \bar{u} \cdot \nabla \underline{u}'_0(N), \quad (12)$$

where  $\underline{u}'_0(n)$  is the value of the turbulent velocity field in the interval  $(n\tau, (n+1)\tau)$  and  $N = t/\tau$  with the usual convention used in numerical coding. We can write for  $\underline{A}$  (the limit  $M \rightarrow \infty$  is understood)

$$\underline{A} = -\frac{1}{M\tau} \sum_{m=0}^{M-1} \int_{m\tau}^{(m+1)\tau} e^{-i\omega t} \underline{f}(t) \cdot \nabla \underline{u}'_0(t) dt. \quad (13)$$

In the interval  $(m\tau, (m+1)\tau)$ ,  $\underline{f}(t)$  is given by Eq. (12) with  $N=m$ . Substituting Eq. (12) into Eq. (13) we find

$$\underline{A} = -\lim_{M \rightarrow \infty} \left[ \left( \frac{e^{i\omega\tau} - 1}{i\omega} \right)^2 \frac{e^{-i\omega t}}{M\tau} \sum_{m=0}^{M-1} \sum_{n=0}^{m-1} e^{i\omega t(n-m)} (\bar{u} \cdot \nabla \underline{u}'_0(n)) \cdot \nabla \underline{u}'_0(m) \right]$$

$$- \lim_{M \rightarrow \infty} \left[ \frac{1}{i\omega} \left\{ 1 + \frac{1}{i\omega\tau} (e^{-i\omega\tau} - 1) \right\} \frac{1}{M} \sum_{m=0}^{M-1} (\bar{u} \cdot \nabla \underline{u}'_0(m)) \cdot \nabla \underline{u}'_0(m) \right]. \quad (14)$$

The first term is a sum over uncorrelated velocity fields  $\underline{u}'_0(n), \underline{u}'_0(m)$

with  $n \neq m$  and, therefore, it vanishes. By the ergodic theorem the sum in the second term is equal to  $\langle \{\bar{u} \cdot \nabla u'_0(m)\} \cdot \nabla u'_0(m) \rangle_E$  where the bracket denotes an ensemble average.

Proceeding in an analogous way with the other terms of  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$  and neglecting the correlations  $\langle \underline{u}'_0 \cdot \nabla p'_0 \rangle_E$ , and  $\langle \underline{u}'_0 \cdot \rho'_0 \rangle_E$  we finally obtain

$$\langle \underline{u}' \cdot \nabla \underline{u}' \rangle = - \frac{e^{i\omega t}}{i\omega} \left[ 1 + \frac{1}{i\omega\tau} (e^{-i\omega\tau} - 1) \right] \langle B \rangle_E. \quad (15a)$$

$$\langle B \rangle_E = \langle \underline{u}'_0 \cdot \nabla(\bar{u} \cdot \nabla \underline{u}'_0 + \underline{u}'_0 \cdot \nabla \bar{u}) + (\bar{u} \cdot \nabla \underline{u}'_0 + \underline{u}'_0 \cdot \nabla \bar{u}) \cdot \nabla \underline{u}'_0 \rangle_E. \quad (15b)$$

It is of interest to consider the limit  $\omega\tau \rightarrow 0$  in Eq. (15). For values of  $t$  such that  $\omega t$  is small, we obtain

$$\langle \underline{u}' \cdot \nabla \underline{u}' \rangle = - \frac{\tau}{2} \langle B \rangle_E \quad (16)$$

This is essentially Elsässer's (1966) expression for  $\langle \underline{u}' \cdot \nabla \underline{u}' \rangle$  valid for steady flows. The above formalism is, therefore, generalized in the present paper for the case of oscillating mean flows. Dropping the subscript 'o' in Eq. (15b), we can write for the  $i$ -th component of  $\langle B \rangle_E$ :

$$\begin{aligned} \langle B_i \rangle_E &= \langle u'_k \partial_k (\bar{u}_j \partial_j u'_i + u'_j \partial_j \bar{u}_i) + (\bar{u}_j \partial_j u'_k + u'_j \partial_j \bar{u}_k) \partial_k u'_i \rangle_E = \\ &\langle 2u'_k (\partial_k \bar{u}_j) (\partial_j u'_i) + \bar{u}_j \partial_j (u'_k \partial_k u'_i) + u'_k \partial_k (u'_j \partial_j \bar{u}_i) \rangle_E. \end{aligned} \quad (17)$$

In Eq. (17) repeated indices are summed and  $\partial_j$  is the derivative with respect to the  $j$ -coordinate. We are particularly interested in studying the effect of the convective motions in the outer solar convection zone on non-radial oscillations. Due to symmetry properties, the correlation coefficients for these motions satisfy certain relations as, for example,  $\langle u'_i u'_j \rangle_E = 0$  unless  $i = j$ ;  $\partial_j \langle u'^2_i \rangle = 0$  unless  $j = r$ . With the help of these relations it is found that

$$\langle 2u'_k (\partial_k \bar{u}_j) (\partial_j u'_i) \rangle_E = \partial_r \langle u'^2_i \rangle_E \partial_i \bar{u}_r, \quad (18a)$$

$$\langle \bar{u}_j \partial_j (u'_k \partial_k u'_i) \rangle_E = \frac{1}{2} \bar{u}_r \partial_r^2 \langle u'^2_i \rangle_E, \quad (18b)$$

$$\langle u'_k \partial_k (u'_j \partial_j \bar{u}_i) \rangle_E = \frac{1}{2} \partial_r \langle u'^2_i \rangle_E \partial_r \bar{u}_i + \langle u'^2_j \rangle_E \partial_j^2 \bar{u}_i. \quad (18c)$$

Therefore,

$$\langle B \rangle_E = (\partial_r v) \bar{v}_r + v \nabla^2 \bar{u} + \frac{1}{2} (\partial_r v) (\partial_r \bar{u}) + \frac{1}{2} \hat{i}_r \bar{u}_r \partial_r^2 v, \quad (19a)$$



where  $\hat{i}_r$  is the unit vector in the radial direction and

$$v = \langle u_i'^2 \rangle_E = \frac{1}{3} \langle \tilde{u}'^2 \rangle_E. \quad (19b)$$

The mean momentum equation (6) can, therefore, be written

$$\frac{\partial \tilde{U}}{\partial t} - \frac{1 + (e^{-i\omega\tau_0} - 1)/i\omega\tau}{i\omega} [(\partial_r v) \nabla U_r + v \nabla^2 \tilde{U} + \frac{1}{2} (\partial_r v)(\partial_r U) + \frac{1}{2} \hat{i}_r U_r \partial_r^2 v] = -\frac{1}{\rho_0} \nabla P + \frac{\Gamma}{\rho_0^2} \nabla p_0 \quad (20a)$$

$$\text{where } \tilde{U} = \bar{u} e^{i\omega t}; P = \bar{p} e^{i\omega t}; \Gamma = \bar{\rho} e^{i\omega t} \quad (20b)$$

### b) The Perturbed Eigenfrequency

Let  $q$  be any physical quantity ( $\tilde{u}$ ,  $p$ , ...). Then  $q = q_0 + Q + q'$  where  $q_0$  is the equilibrium value in the absence of oscillations and turbulence,  $Q (= \bar{q} e^{i\omega t})$  is the oscillatory component and  $q'$  is due to the turbulence (see Eqs. (3)). Clearly,

$$Q = e^{i\omega t} \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T e^{-i\omega t} q dt. \quad (21)$$

We denote by  $\xi$  the displacement due to the oscillations--defined, therefore, by an equation similar to Eq. (21)--and define

$$\delta Q = Q(r_0 + \xi, t) - Q_0(r_0, t). \quad (22)$$

Clearly,  $\partial \xi / \partial t = \delta \tilde{U}$  and, e.g.,  $\delta \nabla U_r = \nabla \delta U_r$ . We apply the  $\delta$ -operator to Eq. (20a) and set  $\omega = \omega_0$  (the unperturbed frequency) in the perturbation terms of Eq. (20a). We obtain

$$-\rho_0 \omega_0^2 \xi - \rho_0 (1 + (e^{-i\omega_0 \tau_0} - 1)/i\omega_0 \tau_0) [(\partial_r v) \nabla \xi_r + v \nabla^2 \xi + \frac{1}{2} (\partial_r v)(\partial_r \xi) + \frac{1}{2} \hat{i}_r \xi_r \partial_r^2 v] = \rho_0 \delta (-\nabla P / \rho_0 + \Gamma \nabla p_0 / \rho_0^2). \quad (23)$$

Lynden-Bell and Ostriker (1967) and Schutz (1979) have developed variational principles to determine the eigenfrequencies of an equation like (23). We define

$$a_\xi = \int \rho_0 \xi^* \xi d^3 r, \quad (24a)$$

$$b_\xi = \int \rho_0 \xi^* (1 + (e^{-i\omega_0 \tau_0} - 1)/i\omega_0 \tau_0) [(\partial_r v) \nabla \xi_r + v \nabla^2 \xi$$

$$+ \frac{1}{2} (\partial_r v)(\partial_r \xi) + \frac{1}{2} \hat{i}_r \xi_r \partial_r^2 v] d^3 r, \quad (24b)$$

$$P_{\xi} = \int \xi^* P \cdot \xi d^3 r, \quad (24c)$$

where  $P$  is given by Eq. (25) of Lynden Bell and Ostriker's paper. Multiplying Eq. (23) by  $\xi^*$  and integrating over the volume, we find

$$-\omega^2 a_{\xi} - b_{\xi} + p_{\xi} = 0. \quad (25)$$

Let

$$\omega = \omega_0 + \omega_1; \quad \xi = \xi_0 + \xi_1, \quad (26)$$

where the subscript '0' denotes now the eigenfrequency and eigenfunction in the absence of the perturbation due to the turbulence. Now for the unperturbed state a variational principle holds; i.e., the quantity  $-\omega_0^2 a_{\xi_0} + p_{\xi_0}$  is of second order in the perturbation. Therefore,  $-2\omega_0 \omega_1 a_{\xi_0} - b_{\xi_0}$  vanishes to first order. That is,

$$\omega_1 = -b_{\xi_0} / 2\omega_0 a_{\xi_0} \quad (27)$$

A word of caution is needed in relation to Eq. (27): It is based on Eq. (19a) for  $\langle \tilde{B} \rangle_E$  and, consequently, in the limit of steady flows the Reynold stresses contain a term,  $-\frac{\tau}{2} \frac{1}{2} \hat{i}_r \bar{u}_r \partial_r^2 v$  which does not contain a derivative of the mean flow. The physical significance of this term is doubtful and, again for steady flows, it can be shown that a more satisfactory theory of "turbulent viscosity" does not include this term. Agreement between the results for oscillatory flows (for  $\tau\omega \rightarrow 0$ ) and steady flows (cf. paper quoted above) is obtained if the Reynolds stresses are given by  $\partial_j(\rho_0 u^j u'^i)$  with

$$\tilde{u}'(t) = \tilde{u}'_0(t) - \int_0^t e^{i\omega t'} \tilde{u}'_0 \cdot \nabla \bar{u} dt'. \quad (28)$$

It can then be readily shown that  $\omega_1 = -b'_{\xi_0} / 2\omega_0 a_{\xi_0}$  with

$$b'_{\xi_0} = (1 + (e^{-i\omega_0 \tau} - 1)/i\omega_0 \tau) I, \quad (29)$$

$$I = \int \xi_0^* \cdot [\partial_r(\rho_0 v) \{ \nabla \xi_{0r} + \partial_r \xi_0 \} + \rho_0 v \{ \nabla \text{div} \xi_0 + \nabla^2 \xi_0 \}] d^3 r;$$

that is,

$$\omega_1 / \omega_0 = F(\omega_0, \tau, a_{\xi_0}) I,$$

$$F(\omega_0, \tau, a_{\xi_0}) = - (1/2\omega_0^2 a_{\xi_0}) [1 - \sin \omega_0 \tau / \omega_0 \tau + i(1 - \cos \omega_0 \tau) / \omega_0 \tau]. \quad (30)$$

## DISCUSSION

It is clear that Eq. (30) allows for the damping or excitation of the oscillations by the turbulent motions as well as for a frequency shift. We notice that the imaginary term in Eq. (30)--namely,  $i(1-\cos\omega_0\tau)/\omega_0\tau$ , is an oscillatory function of  $\omega_0\tau$ . The maximum value occurs at a period of oscillation  $P$ , such that  $P \sim 8\tau/3$ . That is,  $P$  and  $\tau$  are comparable (see Goldreich and Keeley 1977).

To obtain a very rough order-of-magnitude estimate for  $\omega_1$ , we neglect the last two terms in  $I$  and retain only the radial component of  $\xi_0 [= [\eta_{or}, \eta_{oh} \partial/\partial\theta, \eta_{oh} (\sin\theta)^{-1} \partial/\partial\phi] Y_\ell^m(\theta, \phi)]$ . We obtain

$$\omega_1/\omega_0 = F(\omega_0, \tau, a_{\xi_0}) \int \partial_r(\rho_0 v) \partial_r(\eta_{or}^2) d^3r. \quad (31)$$

In Table 1,  $\rho_0 v (=1/3\rho_0 \langle u'^2 \rangle)$  is tabulated for surface values of the solar convection zone.

Table 1. Values of the kinetic energy of the turbulent convective motions as a function of depth in the solar convection zone

d(km)	50	100	200	300	400	500	600
$\rho_0 v/10^3(\text{cgs})$	1.2	7.4	6.4	6.3	6.6	7.1	7.6
d	700	800	900	$10^3$	$10^4$	$10^5$	
$\rho_0 v/10^3$	8.2	8.7	9.3	10.1	$1.3 \cdot 10^2$		$6.8 \cdot 10^2$

It is clear from Eq. (31) that we need to calculate the following expression:

$$A = \int \partial_r(\rho_0 v) \partial_r(\eta_{or}^2) d^3r / (2 \int \rho_0 \eta_{or}^2 d^3r). \quad (32)$$

With the exception of the surface layers of the convection zone, Table 1 shows that  $\partial_r(\rho_0 v) \sim 10^{-4}(\text{cgs})$ . Figure (9) of Ulrich and Rhodes (1983) suggests that the following relations have some validity:

$\rho_0(S) \eta_{or}^2(S) \sim 10 \langle \rho_0 \eta_{or}^2 \rangle$  where  $S$  and the bracket denote a surface and the average value, respectively. Keeping in mind that  $\eta_{or}^2$  decreases rapidly with depth, we find  $A \sim 10^{-4} [\eta_{or}^2(S) \rho_0(S) / \rho_0(S)] / 2d \langle \rho_0 \eta_{or}^2 \rangle$  where  $d$  is the depth of the convection zone. With  $\rho_0(S) = 3 \cdot 10^{-6} \text{ g cm}^{-3}$  (density at a depth of 1000 km),  $A \sim 10^{-8} \text{ s}^{-2}$  and  $\omega_1/\omega_0 \sim 2.5 \cdot 10^{-5}$  for oscillations with a period of 5 minutes. It should finally be noticed that since  $\partial_r(\rho_0 v) > 0$  (a thin surface layer excepted) Equation (31)

shows that the turbulent motions will act as a driving term for the oscillations if  $\partial_r(\eta_{or}^2) > 0$  in the region contributing predominantly to the integral in Eq. (31).

The author is grateful to Drs. F. L. Deubner and J. Leibacher for valuable discussions. The author's work was supported in part by NASA grant W-15.076.

#### REFERENCES

- Baker, N. H. and Gough, D. O. 1979, Ap. J., **234**, 232.  
Deupree, R. G. 1977, Ap. J. **211**, 509.  
Elsässer, K. 1966, Zs. Ap., **63**, 65.  
Gabriel, M., Scuftaire, R., Noels, A., and Boury, A. 1975, Astr. Ap., **40**, 33.  
Goldreich, P., and Keeley, D. A. 1977, Ap. J., **212**, 243.  
Gonczi, G., and Osaki, Y. 1980, Astr. Ap., **84**, 304.  
Gough, D. O. 1977, Ap. J., **214**, 196.  
Keeley, D. A. 1977, Ap. J., **211**, 926.  
Krause, F. 1968, Habilitationschrift, Universität Jena.  
Lynden-Bell, D., and Ostriker, J. P. 1967, M.N.R.A.S., **136**, 293.  
Moffat, H. K. 1978, Magnetic Field Generation in Electrically Conducting Fluids (Cambridge: Cambridge University Press).  
Saio, H. 1980, Ap. J., **240**, 685.  
Schutz, B. F. 1979, Ap. J., **232**, 874.  
Ulrich, R. K. and Rhodes, E. J., Jr. 1983, Ap. J., **265**, 551.  
Unno, W. 1967, P.A.S.J., **19**, 140.  
Xiong, D. 1977, Acta Astr. Sinica, **18**, 86 (translated in Chinese Astr. **2**, 118 (1978)).

# THE STABILITY OF THE LOW DEGREE FIVE MINUTE SOLAR OSCILLATIONS

Russell B. Kidman and Arthur N. Cox  
 Theoretical Division, Los Alamos National Laboratory  
 University of California  
 Los Alamos, NM 87544

In this paper we discuss the decay rate for many of the low degree p modes observed as 5 minute oscillations of the sun. This report is an expanded version of the presentation at Snowmass. These theoretical results use the completely nonadiabatic linear theory of Saio and Cox (1980). Our solar model is based on the evolution results of Christensen-Dalsgaard (1982). Equation of state and opacity data come from the Los Alamos Opacity Library of Huebner, Merts, Magee, and Argo (1977). We compute decay rates for modes ranging from radial ( $\ell=0$ ) to the nonradial ones with  $\ell=5$  for overtones 10 through 28.

Parameters needed for our solar model are given in Table 1. Figure 1 shows the hydrogen mass fraction composition structure. Also given on the figure is the structure given by Christensen-Dalsgaard (1982) for an evolved solar model. Our special equation of state and opacity table with  $X=0.74$  for the hydrogen mass fraction in the outer 0.40 of the mass needs slightly more hydrogen in the central regions than obtained by Christensen-Dalsgaard in order to give a complete and consistent model. The difference in helium production between these two models is about 10%, meaning that the total energy radiated by the sun during its lifetime thus far agrees satisfactorily with accurately calculated evolution sequences.

TABLE 1

SOLAR MODEL:

Luminosity	$3.90 \times 10^{33}$ erg sec <sup>-1</sup>
Mass	$1.989 \times 10^{33}$ g
Radius	$6.955 \times 10^{10}$ cm
Surface temperature	$5.8 \times 10^3$ K
Central temperature	$1.51 \times 10^7$ K
Central density	$122.3$ g cm <sup>-3</sup>
Surface X, Y, Z	0.740, 0.240, 0.020
Central X, Y, Z	0.450, 0.530, 0.020
Depth of convection zone	$0.32R_{\odot}$ ( $0.043M_{\odot}$ )
Temp. at bot. of convec. zone	$2.50 \times 10^6$ K
Central ball	$0.05R_{\odot}$ ( $0.01M_{\odot}$ )
Surface-zone mass	$3.0 \times 10^{22}$ g

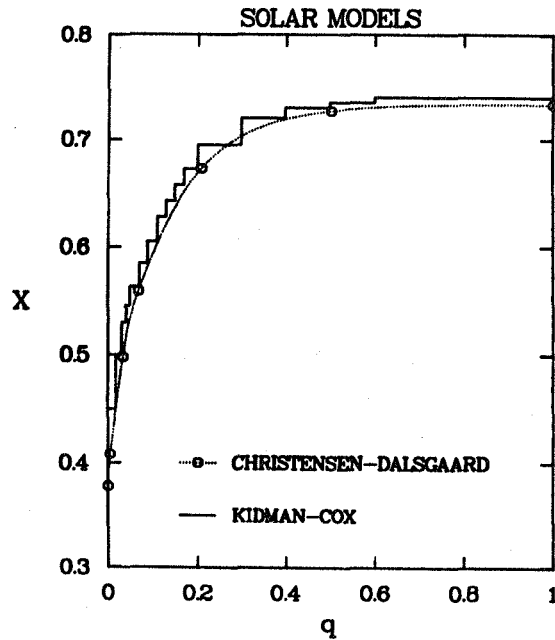


Figure 1. The hydrogen mass fraction of the composition,  $X$ , is plotted versus interior mass fraction of the solar model.

Figure 2 is a plot of the logarithms of temperature (K), opacity ( $\text{cm}^2/\text{g}$ ), and density ( $\text{g}/\text{cm}^3$ ) for our model versus the logarithm of the exterior mass. The very high opacity over the outer 4% of the mass produces a very deep convection zone. The rapid rise of temperature just cooler than 7,000 K requires a small density inversion to give the proper run of pressure to maintain hydrostatic balance. The ratio of mixing length to pressure scale height for all the convection zone is 1.5.

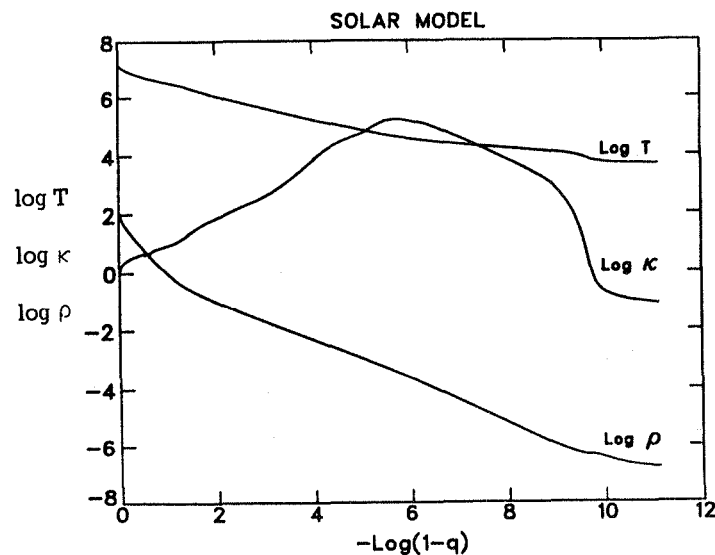


Figure 2. The temperature, opacity, and density structure of the sun is plotted versus the logarithm of the exterior mass fraction.

Solution for the eigenvalues and eigenvectors for the nonradial modes is made for all six of the Dziembowski (1971) variables which have both real and imaginary parts. Figure 3 gives the central variations of  $y_1$ , the Lagrangian variation of the mass shells in the radial direction. The imaginary part, which indicates the variation structure at the mean radius phase of the pulsation as contrasted with the real part applying to the time of the maximum expansion for these linear sinusoidal motions, is very small. This means that the oscillations for this  $p_{23}$ ,  $\ell=2$  mode are very adiabatic. There is little phase change for these lobes which gives essentially standing rather than running waves. Figures 4 and 5 give this same radial variation structure in the outer

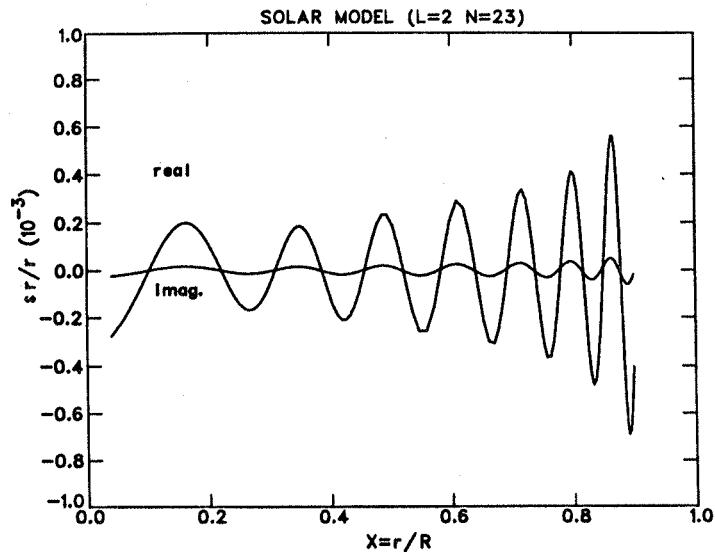


Figure 3. The central variations of the real and imaginary parts of the radial component of the  $p_{23}$ ,  $\ell=2$  oscillations are plotted versus radius fraction. At  $x=0.2$ , 30% of the solar mass is interior. Only 6% of the mass is interior to  $x=0.1$ . At  $x=0.4$  the interior mass is 75% of the model mass.

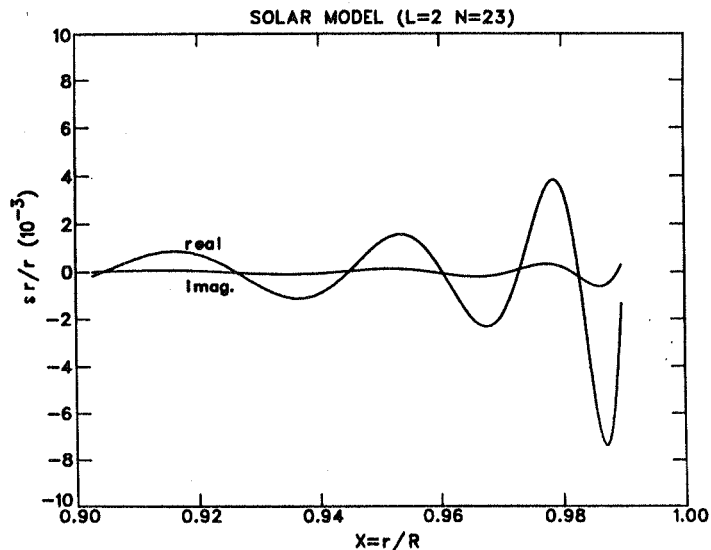


Figure 4. The radial component real and imaginary parts of the  $p_{23}$ ,  $\ell=2$  oscillations are plotted for the outer 10% of the radius.

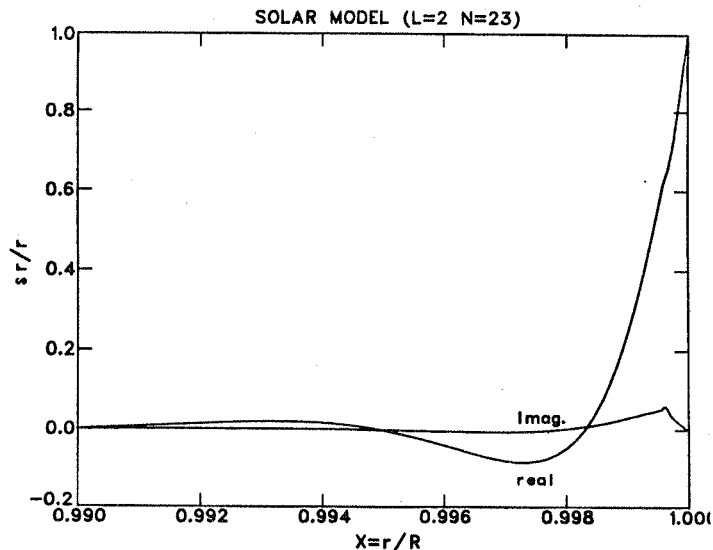


Figure 5. The radial component real and imaginary parts of the  $p_{23}$ ,  $\ell=2$  oscillations are plotted for the outer 1% of the radius.

10% and 1% of the radius. Only at the surface is there any phase change of the real and imaginary parts of this variation, and there also is a significant nonadiabatic effect. At the very surface, the usual normalization gives unity for the real part and zero for the imaginary part. Damping of this mode is related to this feature of the eigensolution, and it is discussed below using the derived complex eigenvalue.

Periods for many modes are given in Tables 2 to 7, respectively, for  $\ell=0$  to 5. Modes  $p_{10}$  to  $p_{28}$  frequencies are listed together with the frequencies observed by Harvey and Duvall (1983). The listed periods and frequencies are from the adiabatic solutions, and the few  $\mu\text{Hz}$  corrections to get the nonadiabatic periods are also given for each mode. Decay rates are calculated by taking the ratio of the imaginary and real parts of the eigenvalue and multiplying this ratio by  $4\pi$ . It is also possible to integrate the P-V loops that all these Lagrangian zones traverse each cycle to get the work done by or done on these zones each cycle. The sum over all the mass zones agrees well with the decay rate derived directly from the eigenvalue component ratio. For each mode the decay times are given in terms of the number of cycles and in terms of the real time.

Our calculations were done with only 329 mass shells and the central ball. These 330 zones are not quite enough to define well the eigensolutions, and we get frequencies 1.3 to 2.5% too large compared with the observed frequencies. See Figure 6 where a comparison of our frequencies with those of Shibahashi et al. (essentially the observed ones) is made. This trend indicates a zoning error, because it decreases with increasing resolution of the pulsation mode and its eigensolutions. Our optimum zoning has very fine mass shells down to a depth of 3.5 million kelvin. In this outer region the temperature increases only about 3% per zone. In the deeper layers, the 100 more zones have a mass ratio from zone to zone of 1.003. We estimate maybe 100 more carefully placed zones could give adequately accurate periods. Our actual



frequencies should not be accepted for high accuracy, but the decay rates listed in the tables should be well determined.

TABLE 2

periods, frequencies and decay rates

l	n	period (sec)	adiabatic frequency (uhz)	non-a incre (uhz)	measured frequency (uhz)	decay by e	
						number of cycles	time (days)
0	10	620.895	1610.6	-15.50	0.0	18330.7	131.73
0	11	573.397	1744.0	-11.97	0.0	8378.6	55.61
0	12	532.832	1876.8	-9.62	0.0	4294.2	26.48
0	13	497.510	2010.0	-8.13	0.0	2481.6	14.29
0	14	466.240	2144.8	-6.83	0.0	1483.9	8.01
0	15	438.540	2280.3	-5.92	0.0	933.4	4.74
0	16	414.175	2414.4	-5.01	0.0	607.0	2.91
0	17	392.092	2550.4	-4.34	0.0	429.3	1.95
0	18	372.124	2687.3	-3.99	0.0	330.4	1.42
0	19	354.083	2824.2	-3.75	0.0	269.5	1.10
0	20	337.481	2963.1	-3.56	0.0	227.6	.89
0	21	322.317	3102.5	-3.41	0.0	191.9	.72
0	22	308.399	3242.6	-3.30	0.0	167.0	.60
0	23	295.583	3383.1	-3.19	0.0	144.7	.50
0	24	283.583	3526.3	-3.16	0.0	132.0	.43
0	25	272.492	3669.8	-3.24	0.0	120.6	.38
0	26	262.244	3813.2	-3.39	0.0	109.4	.33
0	27	252.677	3957.6	-3.48	0.0	105.4	.31
0	28	243.731	4102.9	-3.70	0.0	97.7	.28
*							

TABLE 3

periods, frequencies and decay rates

l	n	period (sec)	adiabatic frequency (uhz)	non-a incre (uhz)	measured frequency (uhz)	decay by e	
						number of cycles	time (days)
1	10	607.745	1645.4	-.58	0.0	13601.4	95.67
1	11	560.942	1782.7	-.59	0.0	6376.5	41.40
1	12	521.033	1919.3	-.54	0.0	3384.0	20.41
1	13	486.203	2056.8	-.54	0.0	1994.7	11.22
1	14	455.839	2193.8	-.58	2161.0	1217.2	6.42
1	15	428.930	2331.4	-.66	2293.0	765.3	3.80
1	16	405.383	2466.8	-.80	2427.0	526.7	2.47
1	17	384.277	2602.3	-.93	0.0	385.9	1.72
1	18	365.017	2739.6	-1.09	0.0	300.8	1.27
1	19	347.465	2878.0	-1.26	2828.0	251.2	1.01
1	20	331.336	3018.1	-1.48	2962.0	211.3	.81
1	21	316.638	3158.2	-1.73	3098.0	180.4	.66
1	22	303.087	3299.4	-2.02	3233.0	157.3	.55
1	23	290.779	3439.0	-2.29	3368.0	142.0	.48
1	24	279.264	3580.8	-2.55	3506.0	128.5	.42
1	25	268.522	3724.1	-2.86	3641.0	116.4	.36
1	26	258.625	3866.6	-3.14	3779.0	108.0	.32
1	27	249.252	4012.0	-3.41	3917.0	102.7	.30
1	28	240.563	4156.9	-3.75	4058.0	95.8	.27
*							

TABLE 4

## periods, frequencies and decay rates

l	n	period (sec)	adiabatic frequency (uhz)	non-a inere (uhz)	measured frequency (uhz)	decay by e	
						number of cycles	time (days)
R	10	588.804	1698.4	.41	0.0	9961.5	67.89
R	11	544.902	1835.2	.34	0.0	4913.6	30.99
R	12	507.211	1971.6	.30	0.0	2740.5	16.09
R	13	473.968	2109.8	.26	2083.0	1620.9	8.89
R	14	444.909	2247.7	.22	2222.0	1009.5	5.20
R	15	419.299	2384.0	.12	2352.0	642.2	3.12
R	16	396.417	2522.6	-.00	2487.0	452.1	2.07
R	17	375.999	2659.6	-.14	2620.0	345.4	1.50
R	18	357.558	2796.7	-.32	2757.0	277.6	1.15
R	19	340.615	2935.9	-.51	2890.0	234.3	.92
R	20	325.108	3075.9	-.74	3024.0	196.9	.74
R	21	310.911	3216.4	-1.00	3161.0	170.4	.61
R	22	297.790	3358.1	-1.31	3295.0	147.1	.51
R	23	285.664	3500.6	-1.57	0.0	135.2	.45
R	24	274.496	3643.0	-1.87	3567.0	123.6	.39
R	25	264.189	3785.2	-1.54	3703.0	111.8	.34
R	26	254.538	3928.7	-2.52	3840.0	107.1	.32
R	27	245.479	4073.7	-2.92	3980.0	99.4	.28
R	28	237.019	4219.1	-3.32	0.0	92.8	.25

TABLE 5

## periods, frequencies and decay rates

l	n	period (sec)	adiabatic frequency (uhz)	non-a inere (uhz)	measured frequency (uhz)	decay by e	
						number of cycles	time (days)
3	10	571.766	1749.0	-.12	0.0	7467.2	49.42
3	11	529.913	1887.1	-.07	0.0	3847.3	23.60
3	12	493.873	2024.8	-.07	0.0	2235.7	12.78
3	13	462.340	2162.9	-.08	0.0	1340.2	7.17
3	14	434.513	2301.4	-.12	0.0	840.4	4.23
3	15	410.225	2437.7	-.18	2408.0	559.8	2.66
3	16	388.319	2575.2	-.29	2542.0	400.9	1.80
3	17	368.460	2714.0	-.42	2677.0	310.6	1.32
3	18	350.527	2852.8	-.56	2812.0	257.3	1.04
3	19	334.074	2993.3	-.74	2948.0	217.1	.84
3	20	319.106	3133.8	-.95	3082.0	184.4	.68
3	21	305.351	3274.0	-1.19	3219.0	161.0	.57
3	22	292.805	3415.2	-1.47	3354.0	142.5	.48
3	23	281.010	3558.5	-1.75	3491.0	129.1	.42
3	24	270.083	3702.6	-2.05	3628.0	117.7	.37
3	25	260.026	3845.8	-2.39	3763.0	108.2	.33
3	26	250.565	3991.0	-2.47	3904.0	103.9	.30
3	27	241.792	4135.8	-3.00	0.0	96.7	.27
3	28	233.600	4280.8	-3.38	0.0	90.9	.25

TABLE 6

## periods, frequencies and decay rates

l	n	period (sec)	decay by e			time (days)	
			adiabatic frequency (uhz)	non-a inre (uhz)	measured frequency (uhz)		number of cycles
4	10	557.147	1794.9	-.32	0.0	5860.3	37.79
4	11	517.054	1934.0	-.25	1918.0	3126.2	18.71
4	12	482.165	2074.0	-.22	2052.0	1838.4	10.26
4	13	451.906	2212.8	-.24	0.0	1125.7	5.89
4	14	425.186	2351.9	-.29	0.0	706.5	3.48
4	15	401.713	2489.3	-.37	2323.0	490.8	2.28
4	16	380.754	2626.4	-.49	2459.0	364.8	1.61
4	17	361.671	2764.9	-.62	0.0	286.6	1.20
4	18	344.215	2905.2	-.77	2728.0	240.6	.96
4	19	328.227	3046.7	-.95	2866.0	201.9	.77
4	20	313.662	3188.1	-1.18	3000.0	173.6	.63
4	21	300.227	3330.8	-1.44	3138.0	150.5	.52
4	22	287.973	3472.5	-1.68	3272.0	137.9	.46
4	23	276.614	3615.1	-1.96	3409.0	125.2	.40
4	24	266.021	3759.1	-2.31	3545.0	112.9	.35
4	25	256.241	3902.6	-2.58	3683.0	106.7	.32
4	26	246.976	4049.0	-2.91	3821.0	99.9	.29
4	27	238.372	4195.1	-3.29	3962.0	93.4	.26
4	28	230.338	4341.4	-3.65	0.0	88.4	.24
x							

TABLE 7

## periods, frequencies and decay rates

l	n	period (sec)	decay by e			time (days)	
			adiabatic frequency (uhz)	non-a inre (uhz)	measured frequency (uhz)		number of cycles
5	10	543.763	1839.0	-.37	0.0	4739.6	29.83
5	11	505.486	1978.3	-.31	1963.0	2621.0	15.33
5	12	471.871	2119.2	-.28	2100.0	1541.7	8.42
5	13	442.601	2259.4	-.30	2235.0	957.2	4.90
5	14	416.968	2398.3	-.36	2371.0	612.0	2.95
5	15	394.088	2537.5	-.45	2505.0	432.6	1.97
5	16	373.676	2676.1	-.58	2641.0	332.0	1.44
5	17	355.271	2814.8	-.74	2777.0	268.6	1.10
5	18	338.341	2955.6	-.89	2914.0	226.2	.89
5	19	322.877	3097.2	-1.10	3050.0	190.1	.71
5	20	308.710	3239.3	-1.32	3187.0	165.1	.59
5	21	295.690	3381.0	-1.61	3324.0	143.3	.49
5	22	283.586	3526.3	-1.85	3460.0	131.4	.43
5	23	272.457	3670.3	-2.14	3597.0	120.3	.38
5	24	262.192	3814.0	-2.49	3733.0	109.3	.33
5	25	252.612	3958.6	-2.73	3871.0	105.2	.31
5	26	243.644	4104.3	-3.11	0.0	97.4	.27
5	27	235.256	4250.7	-3.48	0.0	91.0	.25
5	28	227.276	4399.9	-5.83	0.0	86.0	.23
x							

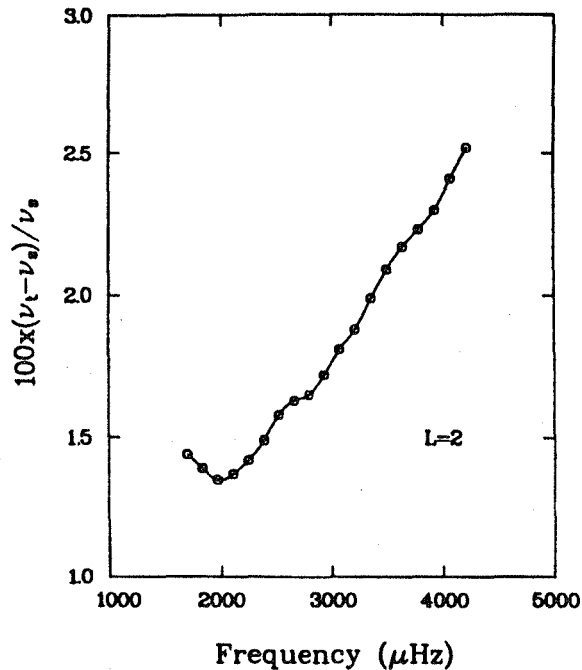


Figure 6. The mode frequency error, as judged from the Shibahashi et al. values, is plotted versus the mode frequency.

A comparison with observations seems to show reasonable agreement with our predictions. The width of the peaks in the power spectrum indicates a decay rate, which is perhaps a matter of days. The coherence of modes over timescales of a month or longer might mean that the mode is reexcited in its existing phase, or it might refer merely to our predicted longer lived modes. Decay rates are faster for higher frequencies because they refer to smaller-scale structures which can more easily gain and lose energy during an oscillation.

Figure 7 shows the work over each pulsation cycle for the outer 30 zones down to a depth of 12,000 K. Actually, the opacity library does not give data for such low temperatures, and the opacities and equation of state are obtained over this region by use of the Stellingwerf (1975ab) analytic fit. The photospheric damping, always assuming a radiation diffusion structure, down to depths of about 40 Rosseland mean opacity mean free paths is the main damping of the solar oscillations. Zone 317 is at  $\tau=10$ , zone 319 is at  $\tau=5$  and zone 324 is at  $\tau=1$ . At least half of the damping occurs where our diffusion approximation is valid. The  $\gamma$  and  $\kappa$  effects of the hydrogen ionization region give some deeper pulsation driving, but it is not enough to self-excite the solar oscillations. The convection zone is neutral because we assume, as is usually done, that the convective flux is frozen at its equilibrium value. Deeper than 10,000 K the very small radiative flux is still modulated, but no more hydrogen or helium driving is significant.

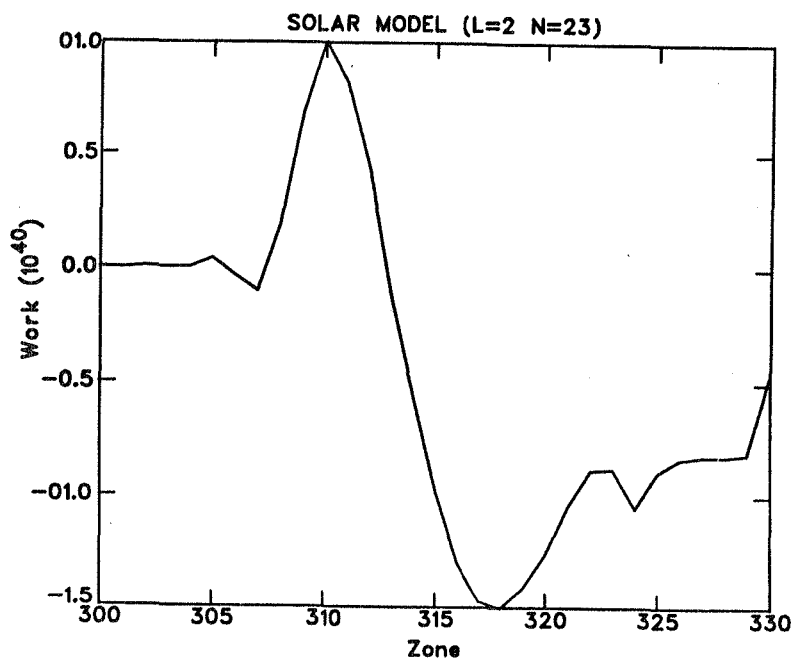


Figure 7. Damping and driving regions of the surface layers are plotted for the outermost 30 zones.

Our conclusions are that all solar 5-minute modes of low order are damped with widely varying rates which increase with frequency. These nonadiabatic effects are confined to the outer  $5 \times 10^{-10}$  of the mass and the outer  $3 \times 10^{-4}$  of the radius.

#### References

- Christensen-Dalsgaard, J. 1982 M.N.R.A.S. 199, 735.  
 Dziembowski, W. 1971 Acta Ast. 21, 289.  
 Harvey, J. and Duvall T. 1983 this conference.  
 Huebner, W.F., Merts, A.L., Magee, N.H., and Argo, M.F. 1977 Los Alamos Scientific Laboratory report LA-6760-M.  
 Saio, H. and Cox, J.P. 1980, Ap.J. 236, 549.  
 Shibahashi, H., Noels, A., and Gabriel, M. 1983, Astron. Ap. 123, 283.  
 Stellingwerf, R.F. 1975a, Ap.J. 195, 441.  
 Stellingwerf, R.F. 1975b, Ap. J. 199, 705.



AN EXCITATION MECHANISM FOR SOLAR FIVE-MINUTE OSCILLATIONS  
OF INTERMEDIATE AND HIGH DEGREE

H. M. Antia, S. M. Chitre and D. Narasimha  
Tata Institute of Fundamental Research  
Homi Bhabha Road, Bombay 400 005  
India

ABSTRACT: The overstability of acoustic modes trapped in the solar convection zone is studied with mechanical and thermal effects of turbulence included, in an approximate manner, through the eddy transport coefficients. Many of these acoustic modes are found to be overstable with the most rapidly growing modes occupying a region centred around 3.2 mHz and spread over a wide range of length-scales. The numerical results turn out to be in reasonable accord with the observed power-spectrum of the five-minute oscillations of intermediate and high degree. The oscillations are probably driven by a simultaneous operation of the  $\kappa$ -mechanism and the turbulent conduction (convective Cowling) mechanism, the dominant contribution to the generation of self-excited acoustic waves arising from the convective Cowling mechanism.

The velocity fields on the solar surface provide a very valuable tool to probe the solar interior. In this connection the discovery of five-minute oscillations by Leighton, Noyes and Simon(1962) has been especially important. The science of solar seismology originated with the observations of Deubner(1975) who resolved the spatial and temporal structure of the sun's five-minute oscillations. The later work of Rhodes, Ulrich and Simon(1977) and Deubner, Ulrich and Rhodes(1979) provided a detailed power-spectrum of the five-minute oscillations of high degree (spherical harmonic degree  $\ell > 150$ ). The observations using integrated sunlight by Claverie et al.(1979) and Grec et al.(1980) revealed the existence of five-minute oscillations of low degree( $\ell < 3$ ). Recently Duvall and Harvey(1983) provided the power spectrum of the five-minute oscillations in the intermediate range( $1 < \ell < 150$ ), and this bridged the gap between the observations of high degree and those of low degree.

The observations have undoubtedly confirmed the suggestion of Ulrich(1970) and Leibacher and Stein(1971) that these oscillations represent non-radial acoustic modes in the solar envelope. The important question of the excitation mechanism responsible for these modes has been examined by various authors. Ando and Osaki(1975) and Ulrich and Rhodes(1977) investigated the stability of non-radial oscillations in a realistic solar envelope model with full effects of radiative heat

exchange. However, the interaction between turbulent convection and oscillations was neglected in these studies. This situation was remedied by Goldreich and Keeley(1977) and Berthomieu et al.(1979) who incorporated the influence of turbulent convection on the stability of acoustic modes to conclude that the turbulent viscosity stabilizes all of these modes.

In the solar envelope, except for the top few hundred Kilometers, major fraction of the total flux is carried by convection; also the turbulent conductivity is much larger than the radiative conductivity for the most part of the convection zone. The turbulence is, therefore, expected to modulate the heat flux in an appreciable manner. This prompted Antia et al.(1982) to undertake a study of the overstability of acoustic modes in the solar envelope with mechanical and thermal effects of turbulence included, albeit in an approximate manner, through the eddy transport coefficients. Many of the acoustic modes trapped in the solar envelope turned out to be overstable with the most unstable modes occupying a region centred around a period of 300 s. Our stability calculations indicated that turbulent heat exchange must play an important role in destabilizing the acoustic modes. Ando and Osaki (1975) had concluded from their analysis that the acoustic modes are largely overstabilized by the  $\kappa$ -mechanism operating in the hydrogen ionization zone, while we recognized that a simultaneous operation of the  $\kappa$ -mechanism and the turbulent conduction mechanism is responsible for exciting the five-minute oscillations. The dominant contribution to the generation of self-excited acoustic waves seemed to arise from the so-called convective Cowling mechanism. It should be emphasised that both the radiative and convective Cowling mechanisms have their origin in the strong superadiabaticity prevailing in the sub-surface layers, but the efficiency of the convective Cowling mechanism is larger by a factor (cf. Unno, 1976)

$$\frac{F^{\text{conv}}}{F^{\text{rad}}} \frac{\nabla}{\nabla - \nabla_{\text{ad}}} .$$

We have been mainly concerned with the stability and driving of the acoustic modes rather than matching the calculated frequencies with observations. Our principal attempt was directed towards the choice of a model which yielded convective modes that are in reasonable accord with the observed features of granulation and supergranulation. Both the convective and acoustic modes are studied adopting the usual linearization procedure and we have included the perturbation in the convective flux in our analysis.

Following Goldreich and Keeley(1977) we have adopted the turbulent viscosity of the form

$$\nu_t = \sigma_t \alpha WL [\min\{ 1, (\frac{1}{\omega t_c})^2 \}] ,$$

where  $\sigma_t$  is the turbulent Prandtl number,  $\alpha$  is an efficiency factor of



order unity appearing in the mixing-length formalism,  $W$  the mean convective velocity,  $L$  the mixing length and  $t_c = L/W$  is the turn-over time for the convective element at that depth and  $\omega$  the frequency of the oscillatory mode in question. The factor in the square brackets ensures that the contribution from only those eddies with turn-over time shorter than  $(1/2\pi)$  times the period of the given mode is included. We have adjusted the value of  $\sigma_t$  so as to obtain the best possible agreement between the length and time-scales of calculated convective modes with the corresponding observed features associated with granulation and supergranulation. This value of  $\sigma_t$  turns out to lie in the range of 0.2 - 0.3 for an envelope model which has a convection zone  $\approx 200,000$  km deep. It is encouraging to find that the same model yields frequencies of acoustic modes which agree reasonably with the observations of Deubner, Ulrich and Rhodes (1979) for high degree ( $\ell$ ) acoustic modes and with Duvall and Harvey (1983) for intermediate values of  $\ell$ . There still remains a marked disagreement with higher harmonics at low  $\ell$ , but this is most probably due to our neglect of the solar interior regions.

We have displayed the results of the stability calculations in Figure 1 which shows the contours of constant stability coefficient  $\eta = (\text{growth rate/frequency})$  of a given acoustic mode in the  $k_h - \omega$  diagram. The outermost contour corresponds to the marginally stable case  $\eta = 0$ , within which all the modes are unstable, while the modes outside the region are stable. We make the plausible assumption that only those modes with significant growth rates will have substantial observed power. It is then interesting to notice that the region where  $\eta > 10^{-4}$  in the Figure approximately coincides with the region where significant amount of power has been observed. In particular, the high frequency cut-off implied by our stability analysis at around 4 - 5 mHz, more or less independent of  $\ell$  is in rough agreement with the observations of Duvall and Harvey (1983). Also, at low  $\ell$ , the lower harmonics are either stable or have an extremely small growth rate, which is consistent with the low observed power in these harmonics.

Our stability analysis bears out the results obtained earlier by Ando and Osaki (1975) that the most unstable p-modes are spread over a region centred mainly around 3.2 mHz with a wide range of horizontal length scales. There is one noticeable difference in our results, namely, we get closed contours of the stability coefficient  $\eta$  with a distinct peak, while Ando and Osaki have open contours with  $\eta$  increasing with  $\ell$  up to  $\ell = 1500$ . This is clearly the influence of turbulent viscosity which because of its effectiveness at small length scales brings the growth rates down sharply at high  $\ell$ .

It is not altogether clear why the amplitudes of these modes are observed to be small. But then it should be remembered that we are dealing with a large number of simultaneously excited modes and we do not know how the modes interact with each other and with the background turbulence in the convection zone.

REFERENCES

- Ando, H. and Osaki, Y.: 1975, Publ. Astron. Soc. Japan 27, 581.  
 Antia, H. M., Chitre, S. M. and Narasimha, D.: 1982, Solar Phys. 77, 303.  
 Berthomieu, G., Cooper, A. J., Gough, D. O., Osaki, Y., Provost, J.,  
 Rocca, A.: 1979 in H. A. Hill and W. A. Dziembowski(eds) "Nonradial  
 and Nonlinear Stellar Pulsations" p307 (Springer, Berlin).  
 Claverie, A., Isaak, G. R., McLeod, C. P., van der Raay, H. B.,  
 Roca Cortes, T.: 1979, Nature 282, 591.  
 Deubner, F. -L.: 1975, Astron. Astrophys. 44, 371.  
 Deubner, F. -L., Ulrich, R. K. and Rhodes, E. J. Jr.: 1979, Astron.  
 Astrophys. 72, 177.  
 Duvall, T. L. Jr. and Harvey, J. W.: 1983, Nature 302, 24.  
 Goldreich, P. and Keeley, D. A.: 1977, Astrophys. J. 211, 934.  
 Grec, G., Fossat, E. and Pomerantz, M.: 1980, Nature 288, 541.  
 Leibacher, J. and Stein, R. F.: 1971, Astrophys. Letters 7, 191.  
 Leighton, R. B., Noyes, R. W. and Simon, G. W.: 1962, Astrophys. J.  
135, 474.  
 Rhodes, E. J. Jr., Ulrich, R. K. and Simon, G. W.: 1977, Astrophys. J.  
218, 901.  
 Ulrich, R. K.: 1970, Astrophys. J. 162, 993.  
 Ulrich, R. K. and Rhodes, E. J. Jr.: 1977, Astrophys. J. 218, 521.  
 Unno, W.: 1976, in E. A. Spiegel and J. P. Zahn(eds) "Problems of Stellar  
 Convection" IAU Colloquium 38, 315.

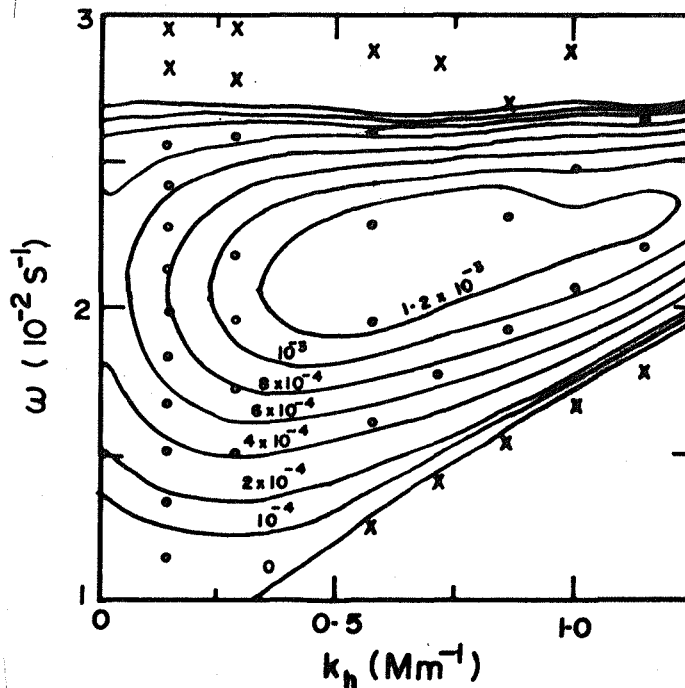


Fig. 1: The contours of equal stability coefficient  $\eta$  are indicated by numbers; the stable modes are labelled by crosses and unstable ones by open circles.

## THE RESONANT COUNT DIAGRAM AND SOLAR g MODE OSCILLATIONS

D. B. Guenther  
Yale University Observatory  
and  
P. Demarque\*  
Goddard Space Flight Center and Applied Research Corporation

Dziembowski (1979) has pointed out that the 160-minute solar oscillation may be the result of resonant three-wave interactions. In a resonant three-wave interaction, two oscillation modes couple together, through the nonlinear terms in the pulsation equations, to drive a beat wave, whose frequency equals the frequency difference of the first two waves. If the beat wave has the same frequency as an oscillation mode of the sun (for example, 104 microHz, the frequency of the 160-minute oscillation), then resonance stimulates the amplitude of this mode. Although this can explain how the g mode oscillations enhance their own amplitudes, so that they can be observed on the surface of the sun, it does not appear to be able to explain why only the 160-minute mode is enhanced. We believe that because of the particular frequency separations of the solar g modes, resonant three-wave interactions stimulate only a selected few g modes. The resonant count diagram provides some evidence for this hypothesis.

The resonant count diagram is obtained by plotting the total number of possible resonant three-wave interactions for a given beat frequency ( $\omega$ ), against the inverse of the beat frequency (the beat period), within a given frequency tolerance  $\Delta\omega$ . The abscissa is the beat period and the ordinate is the total number of interactions such that  $\omega_1 - \omega_2 = \omega \pm \Delta\omega$ . If we assume all the resonant interactions contribute equally to the enhancement of a given beat wave, then the peaks in the curve mark the periods of those modes (beat waves) which are most likely to be significantly enhanced.

We have constructed such a diagram using the  $l = 1, 2, 3, 4$  g modes calculated by Christensen-Dalsgaard, Gough and Morgan (1979) for a standard model of the sun. The diagram has a significant peak at 160 minutes as well as other peaks at longer periods. When we plotted the g modes that Delache and Scherrer (1983) tentatively identified from the Crimea-Stanford data, we found that these modes corresponded with the other peaks in the diagram. This coincidence between the observed g modes and the peaks in the resonant count diagram leads us to believe that the observed g modes do owe their observability to resonant three-wave interactions.

---

\*Permanent Address: Yale University Observatory

If this is the case, the diagram should prove to be a useful diagnostic tool to test the interior mass distribution of solar models. The importance of this diagram can be understood when one considers that the usual method of comparing the theoretical frequency separations of the modes with the observed separations will fail because not all of the modes are enhanced by resonant three-wave interactions. Because the peaks in the diagram correspond to the modes which are most likely to be observed, and because the positions of the peaks depend on the g mode frequency spectrum, which is itself a function of the interior mass distribution of the sun, the diagram can be readily fitted into the observations by adjusting the mass distribution in the solar model. Hence, when a few g modes are positively identified, the diagram should provide a simple method to test our knowledge of the interior physics of the Sun.

See Guenther (1983) and Guenther and Demarque (1983) for a detailed account of this research. The authors wish to thank Bernard Durney and Richard Larson for their useful comments. This research was supported in part by grant AST80-23743 from the National Science Foundation.

#### References

- Christensen-Dalsgaard, J., Gough, D. O, and Morgan J. G. 1979. *Astron. Astrophys.* 73, 121.
- Delache, P. and Scherrer, P. H. 1983. *Nature*, in press.
- Dziembowski, W. 1979 in Nonradial Nonlinear Stellar Pulsation, ed. H. A. Hill and W. A. Dziembowski, p. 272.
- Guenther, D. B. 1983. Ph.D. dissertation, Yale University.
- Guenther, D. B. and Demarque, P. 1983. *Ap.J. Letters*, in press.

FINE STRUCTURE OF SOLAR ACOUSTIC OSCILLATIONS DUE TO ROTATION

Philip R. Goode  
 Arizona Research Laboratories and Department of Physics  
 University of Arizona, Tucson, Arizona, USA

W. Dziembowski  
 N. Copernicus Astronomical Center  
 Polish Academy of Sciences  
 Warsaw, Poland

Ledoux (1951) predicted that rotation causes a fine structure in the spectrum of stellar oscillations. Assuming rigid rotation, he attributed the resultant splitting to linear advection which lifts the degeneracy of each individual oscillation multiplet yielding a spectrum of uniformly spaced fine structure centered about each unperturbed frequency. A more complicated rotation law implies a more complicated spectrum. Thus, the observation of the fine structure of stellar oscillations can be used, in principle, to determine the rotation law inside the star. Our purpose here is to predict the nature of the fine structure of high order, low degree five minute period solar oscillations following from various postulated forms of spherical rotation. We include the first and second order effects of rotation.

The frequency splitting caused by linear advection, assuming spherical rotation, is given by (Hansen, Cox and Van Horn, 1977)

$$\omega_1^{n\ell m} = \frac{\int_0^{R_0} \Omega(r) \left\{ -\left( y_{n\ell}^2(r) + \ell(\ell+1)z_{n\ell}^2(r) \right) + 2y_{n\ell}(r)z_{n\ell}(r) + z_{n\ell}^2(r) \right\} \rho r^4 dr}{\int_0^{R_0} \left( y_{n\ell}^2(r) + \ell(\ell+1)z_{n\ell}^2(r) \right) \rho r^4 dr} \quad (1)$$

where  $\Omega(r)$  is the rotation rate and the normal coordinate is given by

$$\vec{\xi}_{n\ell m} = \left[ \vec{r} y_{n\ell}(r) Y_{\ell}^m(\theta, \phi) + r z_{n\ell}(r) \vec{V}_H Y_{\ell}^m(\theta, \phi) \right] e^{i\omega_0^{n\ell} t} \quad (2)$$

The quantity  $\omega_0^{n\ell}$  is the angular frequency of the oscillation in the absence of rotation. The remaining terms in Equations (1) and (2) have their usual meanings. The expression for the frequency splitting,  $\omega_2^{n\ell m}$ , due to the second order effect of rotation is lengthy. Therefore, we present the asymptotic part of the expression which is valid for acoustic modes (high order and/or high degree),

$$\omega_2^{n\ell m} \approx K_{n\ell}^m \omega_0^{n\ell} \int_0^{R_0} \left\{ \frac{d}{dr}(\epsilon r) y_{n\ell}^2 + \ell(\ell+1) z_{n\ell}^2 \epsilon - \frac{1}{3} \frac{\Omega^2 r}{g} \mu \left[ \left( y_{n\ell}^2 + \ell(\ell+1) \right) z_{n\ell}^2 + \left( 2\mu + \frac{d\ln\mu}{d\ln r} \right) y_{n\ell} \left( z_{n\ell} - y_{n\ell} \frac{g}{\omega_0^2 r} \right) \right] \right\} \rho r^4 dr / \left[ \int_0^{R_0} \left( y_{n\ell}^2 + \ell(\ell+1) z_{n\ell}^2 \right) \rho r^4 dr \right], \quad (3)$$

where  $g$  is the local gravity and

$$\mu = \frac{d\ell n\Omega(r)}{d\ell nr} \quad (4)$$

and

$$K_{n\ell}^m = \frac{\ell(\ell+1) - 3m^2}{4\ell(\ell+1) - 3} \quad (5)$$

where  $K_{n\ell}^m$  represents the result of the angular integration. The quantity  $\epsilon$  is given by

$$\epsilon = \frac{\Omega^2 r}{3g} + \frac{\phi_2}{gr} \quad (6)$$

where  $\phi_2$  is the perturbation of the gravitational potential due to rotation. We wish to thank Douglas Gough (1983) for providing us his equations which enabled us to find our error in the oral presentation. Thus, the results shown here are different from those presented at the conference.

For the purpose of calculation, we assume that the internal solar rotation rate is twice the surface rate from the center to  $R_1$  and smoothly, monotonically decreases to the surface value in going from  $R_1$  to  $R_2$ . We further assume that the rotation rate equals the surface rate from  $R_2$  to the surface (see Figure 1).

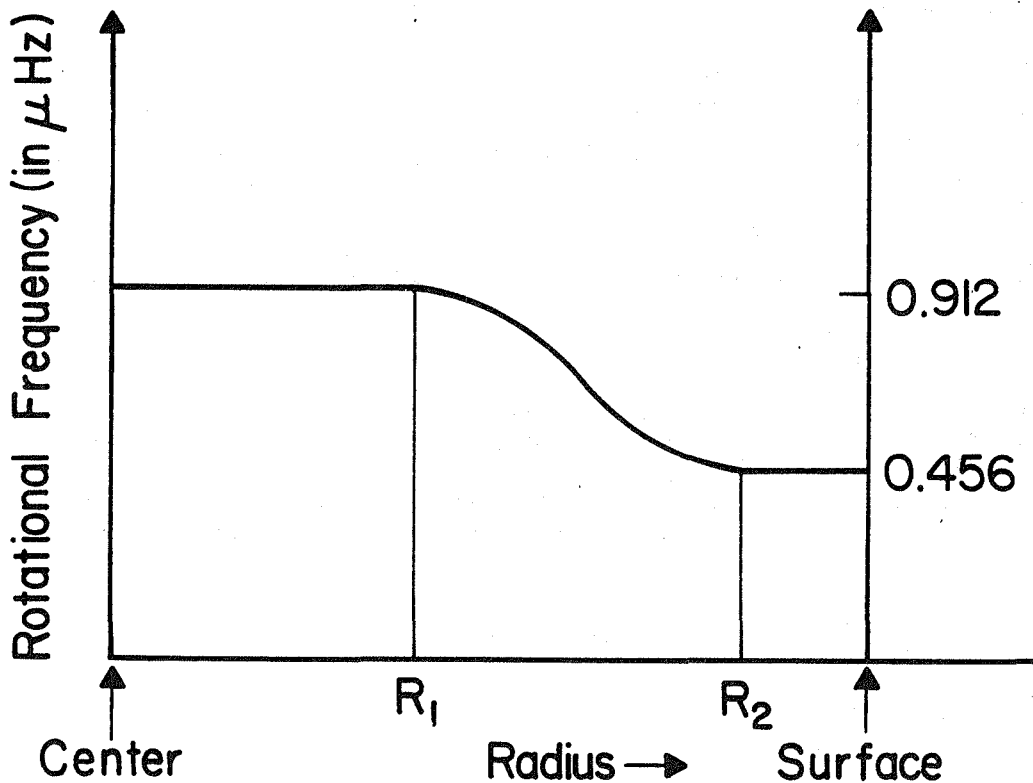


Figure 1. The Internal Rotational Frequency as a Function of Radius.

The predicted splittings which follow from the implementation of the assumed rotation law are given for the  $P_{22} \ell = 1$  five minute period mode for various values of  $R_1$  and  $R_2$  (see Table 1).

Table 1. Predicted First and Second Order Rotational Splitting (in  $\mu\text{Hz}$ ) between the  $m = 0$  and  $1$  components of the  $P_{22} \ell = 1$  mode.

$R_1/R_0$	$R_2/R_0$	$\frac{\omega_1}{2\pi}$	$\frac{\omega_2}{2\pi}$
0.998	0.999	-0.89	1.30
0.990	0.999	-0.89	0.15
0.720	0.999	-0.72	-0.03
0.719	0.720	-0.64	-0.03
	rigid rotation	-0.45	-0.03

The second order effect of rotation is large only if there is a sharp gradient in  $\Omega$  very near the surface. The large effect is a result of the sharp gradient in the region where the splitting kernel is largest. The splitting kernel is the integrand in Equation (1) if  $\Omega(r) = 1$ . For the other cases, including a sharp gradient at the base of the convection zone the second order effect of rotation is small compared to the first order effect of rotation.

One of the authors (P. R. G.) would like to thank the organizers of the conference for financial support which enabled him to attend the conference.

#### References:

- Gough, D. O., private communication, 1983.  
 Ledoux, P. Astrophys. J. 144, 373 (1951).  
 Hansen, C. J., Cox J. P. and Van Horn, H. M. Astrophys. J. 217, 191 (1977).





PART 5

THE SOLAR EQUATION OF STATE



# IONIZATION EQUILIBRIUM AND EQUATION OF STATE IN THE SOLAR INTERIOR

Forrest J. Rogers  
University of California, Lawrence Livermore National  
Laboratory, Livermore, California 94550

## Abstract

Many-body formulations of the equations of state are restated as a set of Saha-like equations. It is shown that the resulting equations are unique and convergent. These equations are similar to the usual Saha equations to the order of the Debye-Hückel theory. Higher order corrections, however, require a more general formulation. It is demonstrated that the positive free energy resulting from the interaction of unscreened particles in high orbits depletes the occupation of these states, without the introduction of shifted energy levels.

## I. Introduction

This paper derives the law of mass action for reacting plasmas from recent many-body quantum statistical results. The primary interest here is to construct a calculational procedure for obtaining the equation of state of the solar interior to high accuracy. Early work on this subject was based on analogy with atomic and molecular gases which have a well defined ideal limit, such that two particle scattering has a negligible effect on the pair distribution function. The earliest work on reacting plasmas is attributed to Saha. The Saha equation follows from the assumptions that the ground state is the dominant contribution to the internal partition function and that the translational partition function is adequately approximated by plane wave eigenstates. Since the internal atomic partition function,

$$Q = \sum_i g_i e^{-\beta E_i}$$

diverges for atoms and ions, the Saha procedure of using only the ground state has frequently been questioned. This has led to the introduction of a number of phenomenological arguments for effectively truncating the sum on  $i$  as a function of plasma conditions.<sup>1</sup> Another problem with the pure Saha procedure is that it, in effect, shuts off the residual Coulomb interaction between atoms and ions. To account for this the Debye-Hückel free energy is usually added to the "ideal" free energy. Higher order non-ideal corrections have also been included.<sup>2</sup> Calculation of the equation of state proceeds by

minimizing a multi-species additive free energy expression, consistent with the model chosen, to obtain the ionization equilibrium. In the present work we loosely refer to all of these free energy approaches as Saha equations. These approaches have the desirable feature that they are computationally simple. However, they have the undesirable features: (1) There is no systematic procedure for including increasingly non-ideal corrections, (2) There are no criteria for determining accuracy.

It is now well established that the divergence of the internal atomic partition function is a non-problem that does not arise in a complete many-body quantum statistical treatment. Its applications to astrophysical problems has been discussed by Ulrich.<sup>3</sup> Even so, the ad hoc method of calculation is still used in almost all astrophysical calculations and may in fact be adequate for many purposes.<sup>4</sup> However, when assurance of accuracy is important the implications of many-body theory cannot be ignored. In the present work we intend to show that, in fact, this leads to a definite procedure which can be systematically improved. It also has associated coupling parameters that indicate its range of validity.

The present method for obtaining the ionization equilibrium and the equation of state of plasmas is somewhat different than usual textbook derivations based on the Saha equation. Consequently, Sec. II shows how the Saha equation can be derived from an activity expansion of the grand canonical partition function. Part A of Sec. III introduces the Beth-Uhlenbeck second virial (cluster) coefficient<sup>5</sup> and discusses how the analytic properties of the two-body trace results in the Planck-Larkin partition function.<sup>6</sup> Part B discusses the resulting, uniquely defined, Saha-type equation that replaces the equation derived in Sec. II. Part C gives an alternative derivation of the Saha-type equation which is obtained from the Helmholtz free energy obtained in Part B. Methods for inclusion of Coulomb corrections are also discussed. Section IV is an extension of the results to two-electron atoms (ions).

## II. Saha Equation Obtained From an Activity Expansion

A number of recent fundamental treatments of ionization equilibrium in plasmas have carried out a renormalization of the grand canonical ensemble.<sup>4,7-10</sup> In the present work it will be useful to illustrate some of the basic aspects of the procedure. To do this we consider the simplest ionization equilibrium, i.e., electrons (e) and nuclei (p) in equilibrium with one electron composite (H):



We begin by assuming that it is possible to carry out the activity expansion of the grand canonical partition function which sequentially switches on two-body, three-body interactions, etc. In a fundamental sense the system described by the Reaction (1) is a two-component sys-

tem of charged elementary particles, some fraction of which are associated in pairs. Truncation of the activity series for this system after two-body interactions are turned on gives<sup>4</sup>

$$\frac{P}{kT} = \frac{1}{V} \ln \Xi = z_e + z_p + z_e^2 b_{ee} + 2 z_e z_p b_{ep} + z_p^2 b_{pp} \quad (2)$$

where  $\rho_i$  is a member of the set  $\{\rho_e, \rho_p\}$

$$z_i = (2s_i + 1) \lambda_i^{-3} e^{\mu_i/kT} \quad (3)$$

$$\lambda_i = (2\pi h^2/m_i kT)^{1/2} \quad (4)$$

is the thermal deBroglie wavelength,  $\mu_i$  is the chemical potential, and the  $b_{ij}$  are cluster coefficients. The activity in Eq. (2) can be eliminated in favor of the density through the relations

$$\rho_i = \frac{N_i}{V} = z_i \frac{\partial (P/kT)}{\partial z_i} \quad (5)$$

Due to mutual repulsion and the large mass of the p-type particles the two-body cluster coefficients for these particles are classical so that,

$$b_{ij} = 2\pi \int_0^\infty r^2 dr (e^{-\beta u_{ij}} - 1), \quad i = j = p \quad (6)$$

where, in this case,  $u_{ij} = u_{pp} = Z^2 e^2/r$  and  $Z$  is the charge of a nucleus of type  $p$ . A similar expression applies to  $b_{ee}$ , but due to the small mass, Wigner-Kirkwood quantum corrections will be required at temperatures that exist outside the solar interior. However,  $b_{ep}$  is fully quantum mechanical and must be calculated from the corresponding analogue of Eq. (6). This is known as the Beth-Uhlenbeck Virial (cluster) coefficient<sup>5</sup> and is given by

$$b_{ep} = b_{ep}^B + b_{ep}^C \quad (7)$$

where  $b_{ep}^B$  and  $b_{ep}^C$  arise from bound states and continuum states, respectively. Explicit expressions for these parts of  $b_{ep}$  are

$$b_{ep}^B = \sqrt{2} \lambda_H^3 \sum_{nl} (2l+1) e^{-\beta E_{nl}} \quad (8)$$

$$b_{ep}^C = \sqrt{2} \lambda_H^3 \left(\frac{1}{\pi}\right) \int_0^\infty \sum_l (2l+1) dp \frac{\partial \delta_l}{\partial p} e^{-p^2/2 R_H kT} \quad (9)$$

where,

$$\lambda_H = (\hbar^2 / 2R_H kT)^{1/2}, \quad (10)$$

$R_H$  is the reduced mass for hydrogen, the  $E_{n\ell}$  are the bound state energies,  $p$  is the relative momentum and  $\delta_\ell$  the phase shift of scattering states having angular momentum  $\ell$ .

In the Saha approximation, since the translational states are plane waves,  $u_{pp} = u_{ee} = \rho_\ell = 0$  so that  $b_{pp} = b_{ee} = b_{ep}^C = 0$ . In addition the divergence of  $Q$  is avoided by truncating the bound state sum of  $b_{ep}^B$  at the ground state, i.e.,  $b_{ep}^B \rightarrow b_{ep}^{gs}$ . Equation (2) now reduces to

$$\frac{P}{kT} = z_e + z_p + 2z_e z_p b_{ep}^{gs} \quad (11)$$

Using the density constraints given by Eq. (5) and the symmetry requirement that  $z_e = Zz_p$  gives the following results (assuming  $Z = 1$ )

$$z_i = [-1 + \sqrt{1 + 8 \rho_i b_{ep}^{gs}}] / 4 b_{ep}^{gs}, \quad i = (e, p) \quad (12)$$

$$\frac{P}{kT} = 2[-1 + \sqrt{1 + 8 \rho_e b_{ep}^{gs}}] / 4b_{ep}^{gs} + [-1 + \sqrt{1 + \rho_e b_{ep}^{gs}}]^2 / 8b_{ep}^{gs} \quad (13)$$

As  $T \rightarrow \infty$   $b_{ep}^{gs} \propto T^{-3/2}$  and as  $T \rightarrow \infty$   $b_{ep}^{gs} \rightarrow \infty$ , so that Eq. (13) predicts that  $P/kT \rightarrow 2\rho_e$  as  $T \rightarrow \infty$  and  $P/kT \rightarrow \rho_e$  as  $T \rightarrow 0$ . These two limits correspond to complete ionization and complete atomization, demonstrating that the activity expression correctly includes the Reaction (1). Since

$$\rho_e = \rho_e^* + \rho_H \text{ and } \rho_p = \rho_p^* + \rho_H,$$

where  $\rho_e^*$  is the number density of free electrons,  $\rho_p^*$  is the number density of free protons, and  $\rho_H$  is the number density of atoms, it follows that  $\rho_e^* = z_e$ ,  $\rho_p^* = z_p$ , and  $\rho_H = 2z_e z_p b_{ep}^{gs}$ . Now Eq. (2) gives

$$\frac{P}{kT} = z_e + z_p + 2z_e z_p b_{ep}^{gs} = z_e + z_p + z_H = \rho_e^* + \rho_p^* + \rho_H, \quad (14)$$

where in the second form of  $P/kT$  we have introduced an activity for atoms,  $z_H = 2z_e z_p b_{ep}^{gs}$ . The result  $z_i = \rho_i$  only results in the ideal limit, at high densities  $z_i < \rho_i$ . We now note that

$$\frac{V z_e z_p}{z_H} = \frac{N_e^* N_p^*}{N_H} = \frac{2V \lambda_H^3}{\lambda_e^3 \lambda_p^3} e^{+E_{1S}/kT} \quad (15)$$

where  $N_i = V \rho_i$ . Equation (15) is the usual Saha equation for hydrogen. This simple example demonstrates how ionization equilibrium arises in the activity series.

### III. A Unique Form of the Saha Equation for Hydrogenic Plasmas

#### A. The Planck-Larkin Partition Function

The derivation of Eq. (15) assumes that  $b_{pp}$ ,  $b_{ee}$  and  $b_{ep}^B \rightarrow 0$  as  $\rho \rightarrow 0$ . It is easy to see that this is not correct. Consider Eq. (6); it is cut-off at small  $r$  by the Landau length,  $\beta e^2$ . Expansion of the exponential in Eq. (6) for the potential  $u_{pp} = e^2/r$  for  $r \geq R$ ,  $u_{pp} = 0$  for  $r > R$  gives

$$b_{pp} \approx 2\pi \int_{\beta e^2}^R r \left( -\frac{\beta e^2}{r} + \frac{(\beta e^2)^2}{2r^2} - \frac{(\beta e^2)^3}{6r^3} \dots \right) dr \quad (16)$$

The first three terms of the expansion  $\rightarrow \infty$  as  $R \rightarrow \infty$ . Similar considerations apply to  $b_{ee}$ . However,  $b_{ep}$  is more complex. In this case  $b_{ep}^B$  and  $b_{ep}^C$  are both divergent.  $b_{ep}^C$  has large distance divergencies similar to  $b_{pp}$  and  $b_{ee}$  involving  $R^2$ ,  $R$ , and  $\log R$  but it has additional divergencies involving  $R^{3/2}$  and  $R^{1/2}$ . The divergencies in  $b_{ep}^B$  also involve half integer powers of  $R$ . This is easily seen by expanding Eq. (1) (in a.u.).

$$Q = \sum_n \frac{1}{n^2} e^{-n^2/T} = \sum_n n^2 + \frac{1}{T} \sum_n \left( \frac{-n^2}{n^2} \right) + \frac{1}{2T^2} \sum_n \frac{1}{n^2} + \dots \quad (17)$$

The two leading terms in the expanded form of Eq. (17) diverge as  $n^3$  and  $n^1$  respectively as  $n \rightarrow \infty$ . For a Coulomb potential the radius,  $r_n$ , of the  $n$ th shell is  $r_n = n^2 a_0$ , so that the fractional power divergencies in  $b_{ep}^B \rightarrow R^{3/2}$  and  $R^{1/2}$  as  $n \rightarrow \infty$ . The fractional power divergencies in  $b_{ep}^B$  and  $b_{ep}^C$ , in fact, compensate leaving only the  $R^2$ ,  $R$ , and  $\log R$  type divergencies that occur in  $b_{ee}$  and  $b_{pp}$ . As a result the effective form of  $b_{ep}^B$  becomes

$$b_{ep}^B = \sqrt{2} \lambda_H^3 Q_{PL} \quad (18)$$

where

$$Q_{PL} = \sum_{nl} (2l + 1) (e^{-\beta E_{nl}} - 1 + \beta E_{nl}) \quad (19)$$

The theoretical basis of Eq. (19) has been thoroughly discussed in the literature<sup>6,8-9,11-12</sup> and is referred to as the Planck-Larkin partition function. The WKB method of Ref. (11) and the fully quantum mechanical method of Ref. (12) demonstrate explicitly how the compensation occurs.

### B. The Saha Equation and the EOS for Low Density Hydrogen

The usual hydrogenic Saha equation was derived in Sec. II. Several unjustified assumptions were made in obtaining this result. We now present the version of this equation that results from many-body theory. Equation (19) has disposed of the divergencies in  $b_{ep}^B$  using only the analytic properties of the two-body partition function. However, the problem of long-range divergencies arising from continuum states remains. The root of the difficulty with these terms can be traced to the activity expansion given by Eq. (2). For a Coulomb potential it is necessary to switch on the full N-body problem at the outset, but only gradually turn on Coulomb coupling by progressively switching on higher powers in the coupling parameter,  $\beta e^2$ . When this is done the sum over similar types of divergencies present in three-body, four-body interactions, etc., yields finite results. One now obtains a pseudoideal gas limit which can be systematically corrected by inclusion of higher order terms in  $\beta e^2$ . In the limit  $\beta \rightarrow 0$ ,<sup>7</sup>

$$\frac{P}{kT} \rightarrow z_e + z_p + \sum_{nl} z_H^{nl} \quad , \quad (20)$$

subject to the conditions

$$\rho_i = z_i \frac{\partial (P/kT)}{\partial z_i}, \text{ where } z_i \text{ ranges over the set } \{z_e, z_p, z_H^{nl}\} \quad , \quad (21)$$

$$z_H^{nl} = 2^{3/2} z_e z_p \lambda_{\bar{H}}^3 \phi_{nl} \quad , \quad (22)$$

$$\phi_{nl} = (2l + 1) (e^{-\beta E_{nl}} - 1 + \beta E_{nl}) \quad , \quad (23)$$

and the  $E_{nl}$  are unscreened hydrogenic energy levels.

Note that Eq. (20) has defined an activity for each bound state. It will be shown that this is necessary for the general problem. For



the present, however, we can define a single activity for atoms according to

$$z_H = \sum_{nl} z_H^{nl} \quad (24)$$

Using Eq. (20-24) the Saha equation that replaces Eq. (5) is

$$\frac{V z_e z_p}{z_H} = \frac{N_e^* N_p^*}{N_H} = \frac{2V \lambda_H^3}{\lambda_e^3 \lambda_p^3 Q_{PL}} \quad (25)$$

Equation (25) is sometimes referred to as the Planck-Larkin equation. The equations of state are given by

$$P_0 = \rho kT, \quad \rho = \sum_i \rho_i \quad (26)$$

$$F_0 = \sum_i F_{i,0}, \quad F_{i,0} \text{ ranges over } \{F_{e,0}^*, F_{p,0}^*, F_{H,0}^{nl}\} \quad (27)$$

$$E_0 = \frac{3}{2} NkT + \frac{N_H}{Q_{PL}} \sum_{nl} (2\ell+1) E_{nl} (e^{-\beta E_{nl}} - 1) \quad (28)$$

where

$$F_{j,0} = -N_j^j kT \ln \left( \frac{2e}{\rho_j \lambda_j^3} \right), \quad F_{j,0} = F_{e,0}^* \text{ or } F_{p,0}^* \quad (29)$$

$$F_{H,0}^{nl} = -N_H^{nl} kT \ln \left( \frac{e \phi_{nl}}{\rho_H \lambda_H^3} \right) \rightarrow F_{H,0} = -N_H kT \ln \left( \frac{e Q_{PL}}{\rho_H \lambda_H^3} \right), \quad N_H = \sum_{nl} N_H^{nl}, \quad (30)$$

$$\rho_H = N_H/V, \text{ and } N = N_e^* + N_p^* + N_H.$$

The second form of Eq. (30), for the total atomic free energy, has assumed the internal states of the atom are occupied according to

$$\frac{N_H^{nl}}{N_H} = \frac{\phi_{nl}}{Q_{PL}} \quad (31)$$

### C. Non-Ideal Saha Equation Derived from the Law of Mass Action

It will be useful to recall the usual method of deriving the Saha equation from the law of mass action. At equilibrium we have from Eq. (28)

$$dF = \frac{\partial F_{e,o}^*}{\partial N_e^*} dN_e^* + \frac{\partial F_{p,o}^*}{\partial N_p^*} dN_p^* + \sum_{nl} \frac{\partial F_{H,o}^{nl}}{\partial N_H^{nl}} dN_H^{nl} = 0 \quad (32)$$

In addition the internal states of the atom (ion) must satisfy the conditions

$$\frac{\partial F_{H,o}^i}{\partial N_H^i} = \frac{\partial F_{H,o}^j}{\partial N_H^j} \rightarrow \ln \left( \frac{\phi_i}{\rho_H^i \lambda_H^3} \right) = \ln \left( \frac{\phi_j}{\rho_H^j \lambda_H^3} \right) \quad (33)$$

where  $i$  and  $j$  correspond to states in the set  $\{nl\}$ . It follows that the ratio of the occupation numbers is given by

$$\frac{N_H^i}{N_H^j} = \frac{\phi_i}{\phi_j} \quad (34)$$

From which it follows that

$$\frac{N_H^i}{N_H} = \frac{N_H^i}{\sum_j N_H^j} = \frac{N_H^i}{[N_H^i + (N_H^i/\phi_i) \sum_{j \neq i} \phi_j]} = \frac{\phi_i}{Q_{PL}} \quad (35)$$

Substitution of Eqs. (33) and (35) into Eq. (32) gives

$$\mu_e + \mu_p = \mu_H \quad (36)$$

where

$$\mu_H = -\ln \left( \frac{Q_{PL}}{\rho_H \lambda_H^3} \right) \quad (37)$$

The Planck-Larkin equation, given by Eq. (25), is again obtained.

In common practice the Debye-Hückel correction is added to  $\beta F_0$ . This is rigorously justified by many-body theory<sup>13</sup> provided  $Q_{PL}$  is used in  $\beta F_0$  and the ratio of the Landau length to the Debye length is small. In this approximation the free energy is given by

$$\beta F = \beta F_0 - V S_R \quad (38)$$

where

$$S_R = \frac{1}{12\pi\lambda_D^3} \quad (39)$$

$$\lambda_D = [kT/4\pi e^2 (\rho_e^* + z^2 \rho_p^* + (z-1)^2 \rho_H^*)]^{1/2} \quad (40)$$

and  $\rho_H = \sum_i N_H^i/V$  ,  $i = \{nl\}$

Equation (38) is a valid approximation when

$$\Lambda = \frac{(\sum_i \xi_i^2 \rho_i) e^2}{\rho kT \lambda_D} \leq 0.1$$

$$\gamma_{ee} = \frac{\lambda_{ee}}{\lambda_D} \leq 0.2 \quad , \quad \lambda_{ee} = \mu / \sqrt{2R_e kT}$$

and  $\alpha_e = \frac{\mu_e}{kT} \leq -1$

where  $\xi_i$  is the net charge on ion  $i$ . Quantum corrections to  $S_R$  are available when high accuracy is required. Since  $(\partial S_R / \partial N_i) = (\partial S_R / \partial N_j)$ , the second form of Eq. (33) still obtains, so that occupation numbers are given by Eq. (35)

The Saha-like equation resulting from Eq. (38) is

$$\frac{N_e^* N_p^*}{N_H} = \frac{2\lambda_H^3 V}{\lambda_e^3 \lambda_p^3} \frac{ze^2/kT\lambda_D}{Q_{PL}} \quad (41)$$

Equation (41) shows that switching on the Debye-Hückel interaction decreases the total number of atoms while, as already shown, the ratios of the occupation numbers are unchanged. In some previous work the Saha equation has been derived for an ideal gas having screened energy levels. Equation (41) gives a precise definition for the analytic form the energy levels should have in this approach, i.e.,

$$\beta E_{nl}^S = \ln \phi_{nl} + ze^2/kT\lambda_D \quad (42)$$

When  $-\beta E_{nl} \gg 1$  Eq. (42) shows that the effective energy levels are given by

$$\beta E_{nl}^S = \beta E_{nl} + ze^2/kT\lambda_D, \quad (43)$$

where the  $E_{nl}$  are unscreened energy levels.

In the opposite limit  $\beta E_{nl}^S \rightarrow \infty$ . When higher order corrections are added to Eq. (41) it is still possible to map the entire ionization equilibrium into an ideal gas form, but involving screened energy levels.

The next non-ideal Coulomb correction beyond Debye-Hückel involves terms that are very similar to second virial (cluster) coefficients for the Debye-Hückel potential. In this approximation

$$\beta F = \beta F_0 - VS_R - VS_2 \quad (44)$$

where

$$S_2 = \rho_e^{*2} s_{ee} + \rho_p^{*2} s_{pp} + 2\rho_e^* \rho_p^* s_{ep}^c \quad (45)$$

$$+ 2\rho_e^* \sum_i \rho_H^i s_{eH}^i + \sum_{ij} \rho_H^i \rho_H^j s_{HH}^{ij}, \quad (45)$$

The  $s_{\alpha\beta}$  are similar to the  $b_{\alpha\beta}$  defined in Eqs. (6-10) except they involve the Debye-Hückel potential and certain low order diagrams are subtracted because they were used in the construction of the screened potential. Also the  $s_{\alpha\beta}$  do not include the Planck-Larkin parts of  $b_{\alpha\beta}$ , since they have already been used in  $\beta F_0$ . Consequently the  $s_{\alpha\beta}$  are well behaved at low T. Methods for calculating the  $s_{\alpha\beta}$  are given in Refs. (8,10,14). The  $s_{\alpha\beta}$  terms involve atoms in different ways, such that for example,

$$\frac{\partial s_{eH}^i}{\partial N_H^i} \neq \frac{\partial s_{eH}^j}{\partial N_H^j} \quad (46)$$

These contributions do not cancel out of Eq. (33). As a result, Eq. (35) is no longer valid and a single chemical potential for atoms cannot be defined. The mass action equations resulting from Eq. (32) now take the form

$$\mu_e^* + \mu_p^* = \mu_H^i \quad \text{for all } \{i\} = \{nl\} \quad (47)$$

The resulting Saha-like equations now involve  $N_H^j$  rather than  $N_H$ . Due to the large core repulsion for ions in excited states it follows that the ratios

$$\frac{N_H^j}{N_H^i} < \frac{e^{-\beta E_j}}{e^{-\beta E_i}}$$

where state  $j$  is a higher lying level than state  $i$ . It was pointed out earlier that the Debye-Hückel interaction increases the state of ionization while leaving the ratios  $N_H^j/N_H^i$  unchanged. The  $S_2$  terms produce both a change in  $N_H$  and the ratios  $N_H^j/N_H^i$ . The change in  $N_H$  is in the opposite direction to the change due to the Debye-Hückel term. Similar considerations apply to higher order terms,  $S_n$ . As a result, the effective energy levels that will map the complete problem into an ideal gas form are very different than Eqs. (42-43) when  $\Lambda > 1$ . It is probably also worth noting that the results being discussed here can be cast in a form such that the low-lying energy levels are screened, but contain only partially the interaction correction. In this case there will be a series of non-ideal corrections. A number of ad hoc approaches have been cast in this form.

It is apparent at this point that a number of different formulations will give the same thermodynamics, but the corresponding energy levels and number densities are somewhat arbitrary. The problem of determining the proper set of number densities required to calculate a particular physical quantity is not taken up in this work.

#### IV. A Unique Saha Equation for Helium-like Ions

The generalization of Eq. (19) to two electron atoms requires a two-electron version of the Planck-Larkin formula. A WKB derivation has been given in Ref. (14). The result obtained is

$$Q_{PL2} = \sum_{n_1 \ell_1} \phi_{n_1 \ell_1} \sum_{n_2 \ell_2} \phi_{n_1 \ell_1 n_2 \ell_2 S} \quad (49)$$

$$n_1 \ell_1 \leq n_2 \ell_2$$

where

$$\phi_{n_1 \ell_1 n_2 \ell_2 S} = g_{n_1 \ell_1 n_2 \ell_2 S} (e^{-\beta E_{n_1 \ell_1 n_2 \ell_2 S}} - 1 + \beta E_{n_1 \ell_1 n_2 \ell_2 S}), \quad (50)$$

the  $E_{n_1 \ell_1 n_2 \ell_2 S}$  are the energies of two electron states relative to one-

electron states, the  $g_{n_1 \ell_1 n_2 \ell_2 S}$  are the corresponding degeneracies, and  $S$  is the spin quantum number. Due to the large separation of the  $1s$  and  $1s^2$  energy levels it is reasonable to assume that doubly excited states make a negligible contribution to the equation of state. In what follows we assume single electron excitations.

Following earlier arguments we obtain the free energy for helium atoms as

$$F_{\text{He},0} = - N_{\text{He}} kT \ln (e Q_{\text{PL2}}^1 / \rho_{\text{He}} \lambda_{\text{He}}^3) \quad (51)$$

where

$$Q_{\text{PL2}}^1 = \phi_{10} \sum_{n_2 \ell_2} \phi_{10} n_2 \ell_2 S$$

The two electron partition function, allowing only single electron excitations, is slightly different than the form suggested by Vorob'ev and Yungman.<sup>15</sup> They demonstrated that the specific heat is sensitive to the exact form of the partition function, so that, the differences could be important in the calculation of  $C_V$  and other derivative quantities. Multiply excited states may also contribute appreciably to derivatives of the equation of state. We now have two reactions occurring

$$\mu_e + \mu_\alpha = \mu_{\text{He}^+} \quad (53)$$

$$\mu_e + \mu_{\text{He}^+} = \mu_{\text{He}} \quad (54)$$

The mass action relation for one-electron ions has already been given. The corresponding result for two electrons composites is

$$\frac{N_e N_{\text{He}^+}}{N_{\text{He}}} = \frac{V \lambda_{\text{He}}^3}{\lambda_e^3 \lambda_{\text{He}^+}^3} \frac{\sum_{n_1 \ell_1} \phi_{n_1 \ell_1}}{Q_{\text{PL2}}^1} \quad (55)$$

Non-ideal corrections can be added similar to the one-electron case. The Debye length that goes in  $S_R$  is now given by

$$\lambda_D = [kT/4\pi e^2 (\rho_e^* + z^2 \rho_\alpha^* + (z-1)^2 \rho_{\text{He}^+} + (z-2)^2 \rho_{\text{He}})]^{1/2}$$

Results similar to Eqs. (30 and 31) are again obtained in the Debye-Hückel approximation. The Saha equation corresponding to Eq. (53) is

modified by a factor  $\text{EXP}(+Ze^2/kT\lambda_D)$  while Eq. (55) is modified by the factor  $\text{EXP}(+(Z-1)e^2/kT\lambda_D)$ .

## V. Conclusion

When Coulomb coupling is weak, e.g. the solar interior, many-body statistical mechanical methods lead to simple procedures for the calculation of ionization equilibrium and the equation of state. These calculations are very similar to the usual Saha-type approaches, but are based on a precise theoretical description of the internal partition function and can be systematically extended to include non-ideal Coulomb interactions. As a result, even though the differences may be small in some instances, there is little reason to introduce phenomenological arguments for weakly coupled plasmas.

## VI. Acknowledgements

This work was performed under the auspices of the U.S. Department of Energy by Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48 and supported in part by the Office of Naval Research.

## VII. References

1. M. McChesney, Canadian J. Phys. 42, 2473 (1964).
2. G. Fontaine, H. C. Graboske, and H. M. Van Horn, Ap. J. Supplement Series 35, 293 (1977).
3. R. K. Ulrich, Ap. J. 258, 404 (1982).
4. F. J. Rogers and H. E. DeWitt, Phys. Rev. A8, 1061, (1973).
5. E. Beth and G. E. Uhlenbeck, Physica 4, 915 (1937).
6. A. I. Larkin, Soviet Phys. JETP 11, 1363 (1960).
7. G. P. Bartsch and W. Ebeling, Beitr Plasma Phys. 11, 393 (1971).
8. W. Ebeling, W. D. Kraft and D. Kremp, Theory of Bound States and Ionization Equilibrium in Plasmas and Solids (Berlin: Akademie-Verlag, 1977).
9. Yu. G. Karsnikov, zh. Eksp. Teor. Fiz. 73, 516 (1977). [Sov. - Phys. - JETP 46, 270 (1977)].
10. F. J. Rogers, Phys. Rev. A24, 1531 (1981).
11. F. J. Rogers, Phys. Lett. 61A, 358 (1977).
12. D. Bolle, Ann. Phys. 121, 131 (1979).

13. H. E. DeWitt, J. Math Phys. 7, 161 (1966).
14. F. J. Rogers, Phys. Rev. A19, 375 (1979).
15. V. S. Vorb'ev and V. S. Yungman, Teplofiz Vys. Temp. 17, 897 (1979).



## THE SENSITIVITY OF SOLAR EIGENFREQUENCIES TO THE TREATMENT OF THE EQUATION OF STATE

**Roger K. Ulrich**

Department of Astronomy, University of California, Los Angeles,  
CA 90024, U.S.A.

**Edward J. Rhodes Jr.**

Department of Astronomy and Earth and Space Science Institute,  
University of Southern California, Los Angeles, CA 90007 and  
Jet Propulsion Laboratory, Pasadena, CA 91109, California Institute  
of Technology, U.S.A.

**Abstract:** We have examined the sensitivity of solar eigenfrequencies to uncertainties in the equation of state. The principal uncertainties in the equation of state involve the treatment of pressure ionization, the Debye-Huckel coulomb corrections and the treatment of many-particle interaction effects. We find that for the lowest degree modes ( $l$  between 0 and 3) the terms and procedures used in our equation of state which deal with these uncertainties introduce changes in the frequencies which are less than  $4 \mu\text{Hz}$ . Recently, Shibahashi, Noels and Gabriel (1983) published solar eigenfrequencies using a theory with an equation of state improved with respect to the theory used earlier by Shibahashi and Osaki (1981). Their comparison between the two sets of results suggested that uncertainties in the frequencies as large as  $10 \mu\text{Hz}$  could be caused by the equation of state. We feel that since the entire effect of the uncertain terms is only  $4 \mu\text{Hz}$  and since the uncertainties are only a fraction of each term, the  $10 \mu\text{Hz}$  changes found by Shibahashi et al must be a consequence of differences between the earlier and later calculations in areas other than the equation of state.

### 1. Introduction

Although the solar equation of state is largely uncomplicated, there remain uncertainties in the areas of pressure ionization, free particle scattering states and the effects of multi-particle interactions. The magnitude of the potential problem is made apparent by a simple application of Saha's equation to the solar center which shows that without any additional physical effects beyond simple ionization roughly 25% of the hydrogen should be neutral even though the temperature is 15 million degrees. In fact additional physical principles come into the problem and reduce the abundance of incompletely ionized species well below the level indicated by the Saha equation. A discussion of the proper approach to the equation of state is given elsewhere in this volume by Rogers (1984). We give here a very brief summary of the areas of uncertainty.

First, the effect due to the electrostatic interaction between the charged particles alters the pressure and energy density in a manner which is frequently treated by the Debye-Huckel approximation. The term which is added to the thermodynamic quantities due to this approximation is sometimes called the

Coulomb correction. The presence of the plasma charge surrounding each nucleus produces a screening effect which alters the bound electron energy levels so that they approach the continuum level and eventually the electron becomes unbound. Second, a cluster expansion (see Mayer 1958 and Eyring, Henderson, Stover and Eyring 1982 for discussions of the cluster expansion method) of the activity (the thermodynamic quantity which is like the pressure but which is most readily calculated from the density) in terms of the density shows that the bound states and the states of free electrons near nuclei (called scattering states) produce counteracting effects on the activity. To lowest order the combined result can be represented by using an effective internal partition function in the ionization equilibrium equation which replaces the exponential factors of the form  $\exp(E/kT)$  in the Saha equation with the Planck-Larkin factor  $[\exp(E/kT) - 1 - E/kT]$ . This factor has an asymptotic value of zero instead of unity when  $E/kT$  approaches zero. Third, the complications from many-body interactions are not eliminated by this form of the Saha equation; instead their effects are shifted from the formation of bound states to the alteration of higher order cluster terms in the description of the scattering states. Ebeling and Sandig (1973) have evaluated these terms for a gas of pure hydrogen and given their result in the form of a Pade approximation to the leading additional term in the cluster expansion - the 2nd Virial term.

We present here eigenfrequencies for solar models which were calculated to study the effects of the individual terms in the equation of state. These models were originally presented in the paper by Ulrich (1982). The treatment of the changes in the energy levels due to screening in that paper is not optimal according to the discussion by Rogers in this volume. Nonetheless, the effect of partial ionization was found to be entirely negligible as long as the Planck-Larkin form of the ionization equilibrium equation was used. The recent discussion by Rouse (1983) of the Planck-Larkin equation is based on the faulty assumption that the effective partition function is the sum of actual occupation numbers of the available energy levels. Although for a non-interacting gas this would be a valid assumption, in the case of the solar equation of state it is not. The compensation of the effects of the bound states by the scattering states represented by the Planck-Larkin equation results from the reduction of the number of unbound states available to free electrons because the wave functions of the scattering states are distorted by the presence of the nuclear charge. Rogers (1977) has given a discussion of this compensation in terms of a WKB description of the wave functions. In order to make comparisons with astrophysical data, the actual atomic occupation numbers must be calculated separately from a thermodynamically self-consistent formulation. Such a calculation has not yet been made.

## 2. The Models

The models which we use here were all computed within a few days of each other. During this period, the only quantities which were varied were the various equation of state terms. We have subsequently made other changes and improvements which have made the 'Standard Model' of the Ulrich (1982) paper obsolete. Nonetheless, this set of models is valuable because it provides an opportunity to study the equation of state effects in isolation. Because the Planck-Larkin form of the ionization equilibrium equation is well established, we have concentrated on the three models of the Ulrich (1982) paper which use

that formulation to treat the ionization explicitly instead of artificially assuming full ionization. The three models are Model 19 which omits both the Coulomb correction and the scattering state collective effects, Model 21 which includes the Coulomb correction but omits the collective effects, and Model 22 which includes both effects. We find that the frequencies change by less than 4  $\mu\text{Hz}$  as a result of any of the equation of state terms. For comparison, we also computed frequencies for model 10 which incorrectly uses the Saha equation throughout the solar interior. This model includes such erroneous results as an abundance of neutral hydrogen at the solar center equal to 0.25 even though there would be inadequate volume for so many bound electrons without having adjacent wave functions overlap.

### 3. Theoretical Frequencies

We have computed frequencies of the solar oscillations using the methods described by Ulrich and Rhodes (1983). We adopt here a slightly different

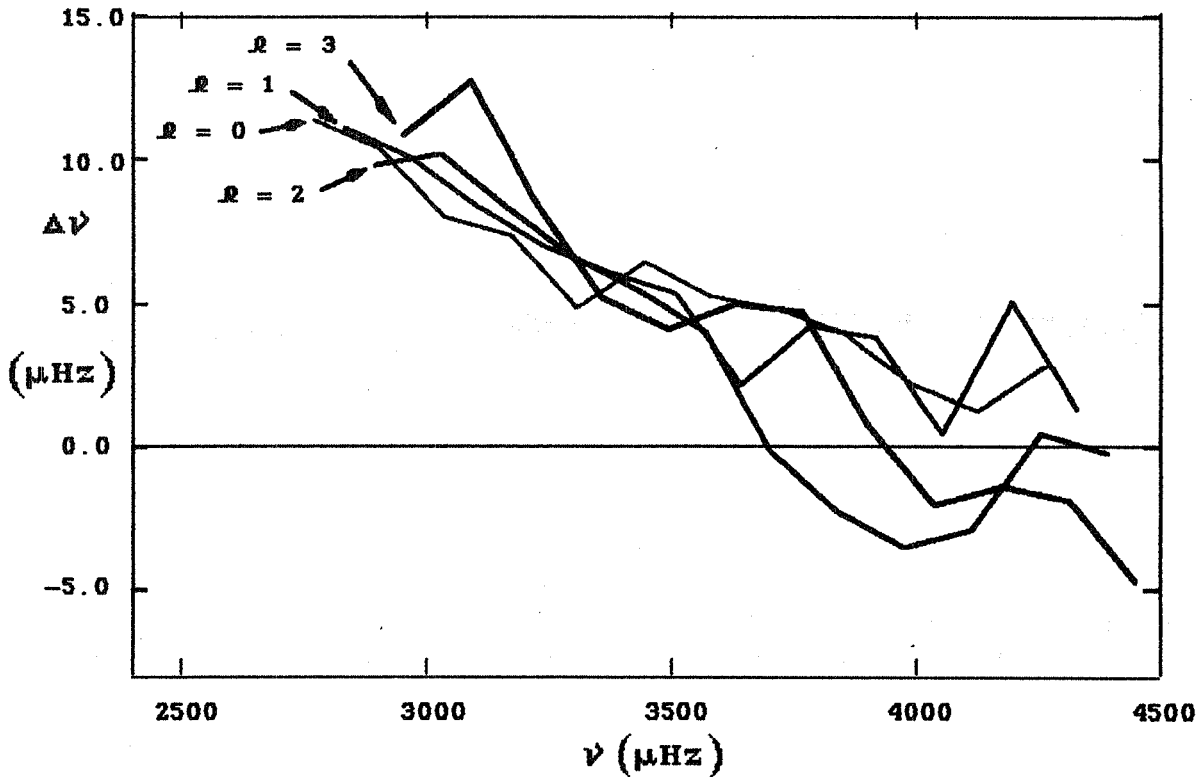


Figure 1 Comparison between the theoretical and observed frequencies. The plot shows the frequency differences: observed minus theoretical. The observations are from Grec, Fossat and Pomerantz (1983) and the theory is for the standard model which incorporates the Planck-Larkin ionization equation, the Coulomb corrections and the 2nd Virial term. The values of  $\gamma$  are indicated by the thickness of the lines with the thinnest having  $\gamma = 0$  and the thickest having  $\gamma = 3$ .

format for the presentation of the results than we have used in the past. Rather than showing the individual frequencies on some form of folded plot, we show only differences between frequencies. Figure 1 shows the differences between the observations and the frequencies for the standard model. Each line in this figure shows the discrepancy as a function of the frequency for fixed values of  $\lambda$ . The erratic behavior at the higher frequencies in this figure and subsequent figures is caused by the presence in the theory of a resonant region in the chromosphere. When the interior mode frequencies and the chromospheric mode frequencies are nearly equal then the chromospheric mode can perturb the interior mode frequency. The existence of this perturbation enhances the sensitivity of the frequencies to changes in the models.

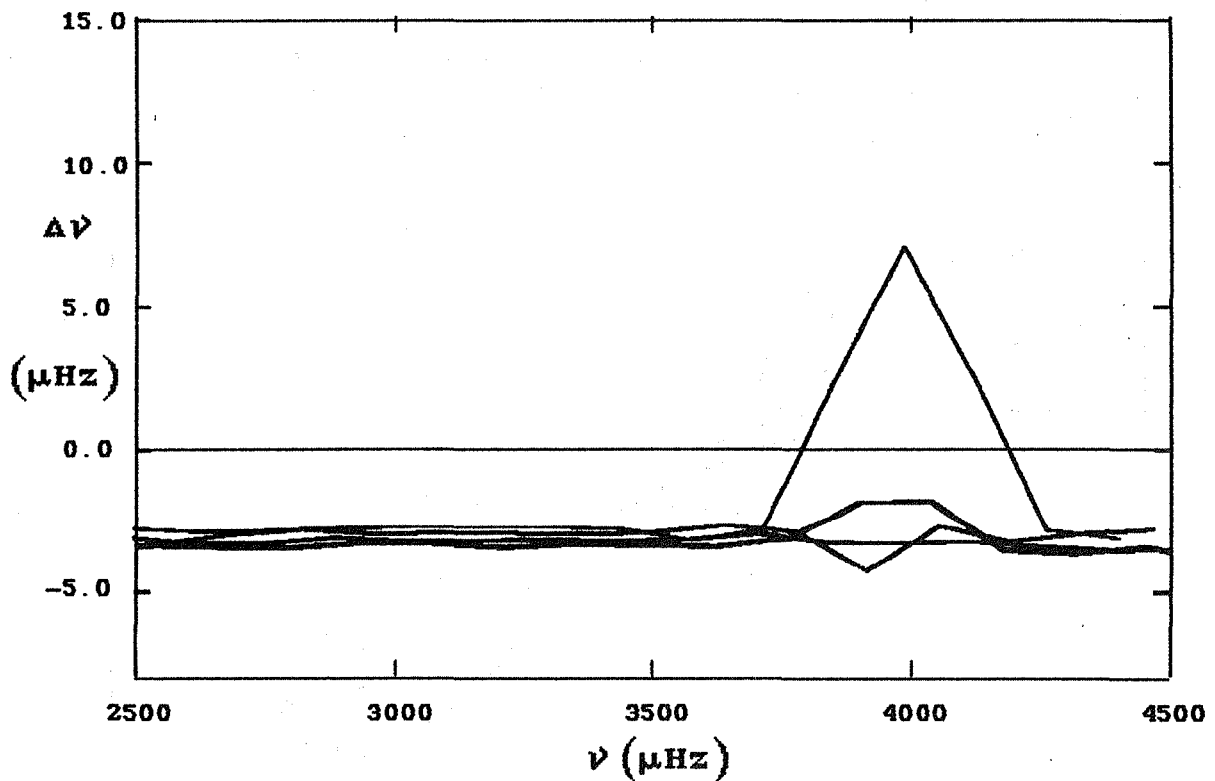


Figure 2 The effect of adding the collective effects as represented by the 2nd Virial term. Frequency differences for the model with the term minus the model without the term are plotted. The values of  $\lambda$  are indicated as in Figure 1.

Figures 2 and 3 show the changes in the oscillation frequencies due to the influence of the collective effects as treated in the form of the 2nd virial term and the influence of the Coulomb corrections. The frequency dependence of the effects are nearly independent of  $\lambda$  as long as  $\lambda$  is near 0. The eigenfunctions for these oscillations have very similar dependence on radius in the outer layers of the sun. Since the largest changes in the equation of state due to the collective effects and the Coulomb corrections occur in the outer

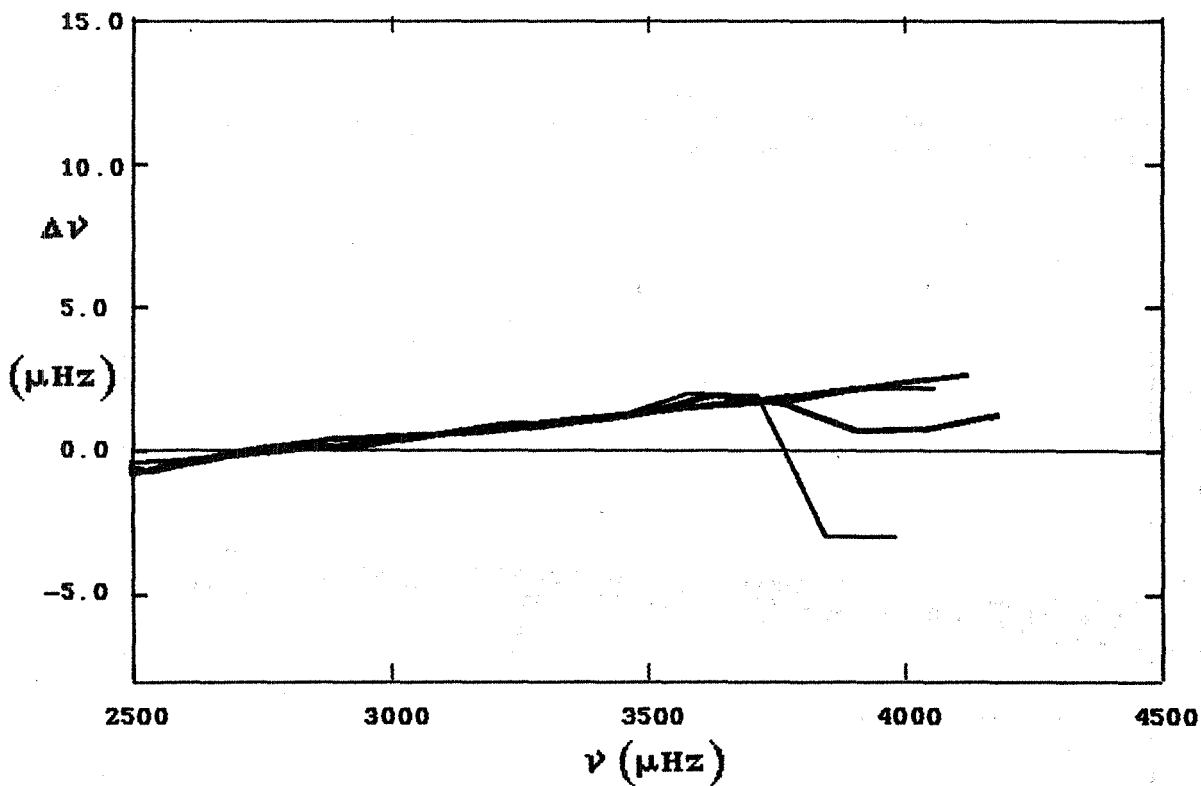


Figure 3 The effect of adding the Coulomb correction. Frequency differences for the model with the correction minus the model without the correction are plotted. The values of  $\lambda$  are indicated as in Figure 1.

sections of the sun, the actual value of  $\lambda$  has little impact on the frequencies unless  $\lambda$  is very much larger than 4. Both the magnitude and dependence on  $\nu$  of the changes due to the coulomb corrections and the collective effects are inadequate to play a major role in explaining the discrepancy between theory and observation. One other equation of state procedure is of interest even though it is not really an area of uncertainty and that is the effect of using the Saha equation throughout the solar model. The results of this calculation are not to be taken seriously since the high abundance of neutral hydrogen predicted for the solar core with this equation is unphysical due to the scattering states and the excessive volume occupied by the bound electrons. Nonetheless, this equation of state is significantly different from the one used in the standard model and therefore gives some indication of an upper bound on the range of variation in the frequencies which might be caused by a major change in the equation of state. Figure 4 shows a comparison between two models which differ only in their treatment of the ionization equilibrium. Both models included the Coulomb corrections but the model using the Planck-Larkin equation omitted the 2nd Virial Term. The frequency changes shown in Figure 4 are larger than those in Figures 2 and 3 but still not large enough to explain the discrepancy between the theory and observation and the model with the Planck-Larkin equation is in slightly better agreement with observation than the model with the Saha equation. Note the wavy nature of the frequency changes. This pattern is caused by the fact that the largest changes between the two models are localized near

the model center. As the eigenmode structures change with frequency, maxima and minima in the eigenfunctions move past this localized region and cause the frequency shifts to vary in response.

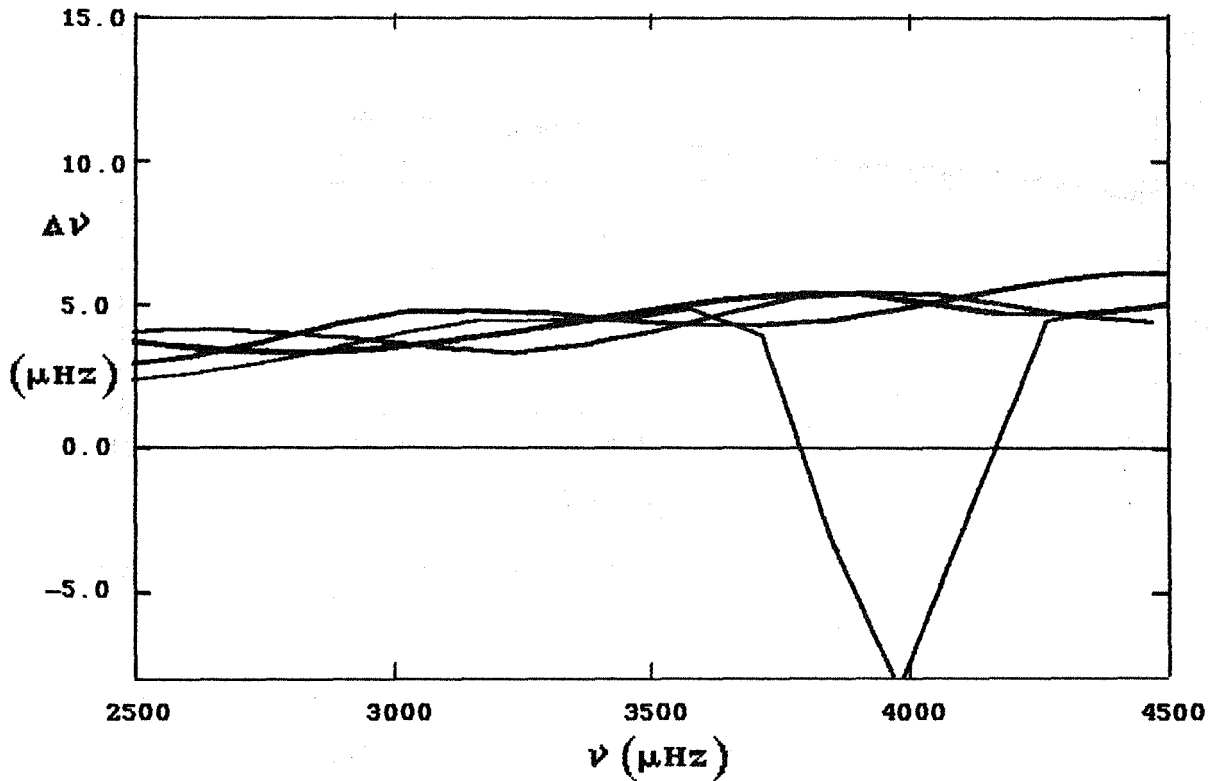


Figure 4 The effect of using the Saha equation for the ionization equilibrium instead of the Planck-Larkin equation. Frequency differences for the model with the Planck-Larkin equation minus the model with the Saha equation are plotted. The values of  $\lambda$  are indicated as in Figure 1.

Recently, Shibahashi, Noels and Gabriel (1983) have studied the influence of the equation of state on the solar eigenfrequencies and state that the inclusion of the coulomb corrections improves the agreement between theory and observation. Our calculations which explicitly study the influence of this term are in disagreement with their conclusion. We feel that the changes between the Shibahashi and Osaki (1981) frequencies and the Shibahashi, Noels and Gabriel (1983) frequencies are most probably a consequence of other modifications in their numerical procedures such as possibly the interpolation in the Los Alamos Opacity Library.

## References

- Ebeling, W. and Sandig, R. 1973, *Ann. Physik*, 28, 289.
- Eyring, H., Henderson, D., Stover, B. J., and Eyring, E. M. 1982, **Statical Mechanics and Dynamics**, John Wiley & Sons: New York, p. 548-584.
- Greig, G., Fossat, E. and Pomerantz, M. 1983, *Solar Physics*, 82, 55.
- Mayer, J. E. 1958, **Handbuch der Physik**, Springer, Berlin, 12, 123.
- Rogers, F. J. 1977, *Physics Letters*, 61A, 358.
- Rogers, F. J. 1984, This Volume.
- Rouse, C. A. 1983, *Ap. J.*, 272, 377.
- Shibahashi, H., Noels, A. and Gabriel, M. 1983, *Astron. and Astroph.*, 123, 283.
- Shibahashi, H., and Osaki, Y. 1981, *Publ. Astron. Soc. Japan*, 33, 713.
- Ulrich, R. K. 1982, *Ap. J.*, 258, 404.
- Ulrich, R. K. and Rhodes, E. J., Jr. 1983, *Ap. J.*, 265, 551.

**End of Document**