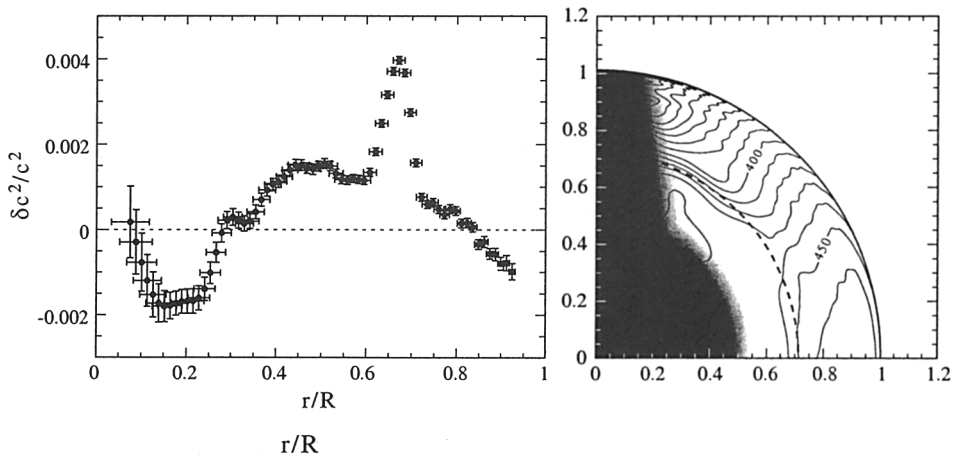


Session I

Global Structure and Evolution of the Solar Interior



The relative difference between the spherically averaged sound speed of the sun and the sound speed of a standard solar model (left panel). The right panel shows a contour plot of the sun's angular velocity (adapted from Gough, p. 3).

The New Era in Helioseismology

D. O. Gough

*Institute of Astronomy and Department of Applied Mathematics and
Theoretical Physics, University of Cambridge, CB3 0HA, England*

Abstract. The end of the millennium marks the beginning of the third phase of helioseismology. The first phase was the establishment of the initial astronomical inferences, such as estimates of the depth of the solar convection zone and the protosolar helium abundance, obtained by comparing the seismic properties of theoretical solar models with the first wave of helioseismic data acquired using instruments that had not been designed for the purpose. The second phase was the determination of the spherically symmetric component of the hydrostatic stratification throughout most of the solar interior, and the angular velocity, using inverse methods to analyse the frequencies of normal modes estimated from data obtained most recently from purpose-built networks of ground-based observatories and from space. We have reached the point beyond which further pursuit of the now-well-tried methods to improve the inferences will be apparently slow. The next era will be characterized by painstaking attention to detail, to extract a new level of precision necessary to isolate subtle properties of the sun for asking more sophisticated questions. We are already seeing the normal-mode representation of helioseismic waves being complemented by other representations that may be more suitable for investigating inhomogeneity and time variability particularly of the sun's surface layers. The outcome will enable us to address more accurately issues concerning global dynamics, the equation of state and the chemical composition, and also the properties of convection and the seat of solar activity.

1. Introduction

When discussing the physics of solar oscillations I shall adopt the normal-mode/wave-propagation duality, recognizing that true linearized normal modes of an idealized steady star can be represented as a superposition of propagating waves, and vice versa. Normal modes have been used mainly to infer the global stratification and rotation of the sun; the constituent waves are being used mainly for inferring inhomogeneity.

As a foundation for discussing where the subject is heading I present in Figure 1 two well-known plots: the first is the relative difference between the spherically averaged sound speed of the sun and the sound speed of a standard (therefore, spherically symmetric) solar model; the second is a contour plot of the sun's angular velocity. To have inferred these quantities in the interior of

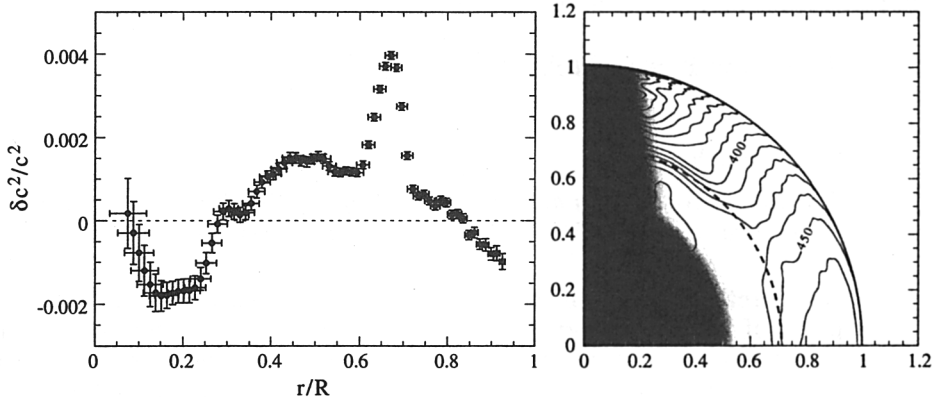


Figure 1. The left-hand panel depicts the relative difference $\delta c^2/c^2$ between the spherical average of the square of the sound speed in the sun and the square of the sound speed in the standard solar model S of Christensen-Dalsgaard *et al.* (1996). The points plotted are weighted averages, with weight functions whose characteristic widths (roughly, the full width at half-maximum) are denoted by the horizontal bars; the vertical bars represent \pm one standard deviation of the uncertainties in the values of the averages resulting from the formal errors in the measured oscillation frequencies, assuming the latter to be uncorrelated. The errors in the sound-speed averages plotted, however, are correlated (cf. Fig. 4). The hump centred at $r/R \approx 0.66$ is not a hump in the sound speed in the sun, but is the outcome of a localized smoothing of a sharp decline with depth (relative to a relatively smoothly increasing function) immediately beneath the convection zone of the magnitude of the slope of c^2 in the model, that decline having been produced by an augmentation of the helium abundance resulting from gravitational settling. The smoothing is probably a result of material redistribution by macroscopic motion in the tachocline. The inversions are of SOI/MDI frequency data (from Kosovichev *et al.*, 1997a). The right-hand panel depicts contours of constant north-south averaged angular velocity Ω in a quadrant of a plane through the rotation axis of the sun. The averages are optimally localized (subject to a tradeoff between degree of localization and susceptibility to data errors). The contour with $\Omega/2\pi = 400 - 450$ nHz are labelled; the others are drawn at intervals of 10 nHz. The blackened region is where the data provide no trustworthy information (from Schou *et al.*, 1998).

a star is a major triumph of helioseismology. The outcome has had a profound impact on heliophysics, eliminating most proposals for alleviating the neutrino problem and constraining dramatically our view of solar spindown. In addition, our knowledge of the stratification of the sun has improved our description of the microphysical properties of stellar material. In particular, it has revealed an error in the computations of the tables of stellar opacities that had been supplied to astrophysicists, the correction of which subsequently removed several theoretical problems concerning stellar variability for which that error was responsible; moreover, it has provided information about the adiabatic compressibility of the (dense) plasma in the solar convection zone which is precise enough to estimate the ability of neighbouring ions and electrons to destroy coherent excited states of neutral hydrogen atoms and neutral and singly ionized helium.

Both of the inferences illustrated in Figure 1 were drawn using a normal-mode representation of solar oscillations. Wave representations have been used to confirm these results and, more usefully, to infer lateral inhomogeneity asso-

ciated with convection and magnetic activity. I shall discuss such issues, quite briefly, later. First, I shall address some of the improvements that will be required in order to extend the usefulness of our inferences concerning large-scale structure.

2. The solar core

Despite the evidence that, broadly speaking, the deficiency of observed solar neutrinos is the result of transitions of electron neutrinos to other flavours, it is still important to neutrino physics that we constrain the structure of the solar core more tightly. The reason is that the evidence from SuperKamiokande (Kajita, 1999) is that the electron-neutrino decay length is greater than the diameter of the earth, which leaves the sun as being still the most important neutrino source. It is important, therefore, to understand that source.

It is interesting to study the core not only for addressing neutrino physics, but also because there is an important property of the core that is widely ignored by stellar physicists: namely that it is linearly unstable to *g* modes. What the nonlinear development of that instability is is open to debate, although current estimates suggest quenching at a low amplitude by stable coupled *g* modes of higher order and degree (Dziembowski, 1983; Jordinson and Gough, 2001). Nevertheless, a certain degree of large-scale material redistribution is not out of the question.

In Figure 2 I show a plot of the spherical average of the square of the sound speed c in the core inferred from a combination of low-degree VIRGO data and intermediate-degree MDI data; it is compared with Christensen-Dalsgaard's (1996) standard model S and a model by Saio. Although there is a difference between the absolute values of c^2 of the two theoretical models, both have similar curvature. The curvature is determined by the derivatives of the opacity κ and the nuclear energy generation rate ϵ with respect to density, temperature and chemical composition (which are more robust than the absolute values of κ and ϵ), and by the radial distribution of the chemical abundances, principally the helium abundance Y (or, almost equivalently, the hydrogen abundance X). The positive gradient of c^2 near the centre of the sun is due to the negative gradient of Y , which is brought about by more hydrogen having been converted into helium in the innermost regions. The apparently gentler variation of the solar c^2 , if it is correct, is perhaps a symptom of an error in modelling. One might entertain the possibility that less inhomogeneity has been established by the nuclear reactions, perhaps because the sun is somewhat younger than the model — the amount by which it would need to be younger is not wholly implausible. But alternatively there may have been a slight homogenization, due to motion which may still be occurring today, reducing the variation in Y and perhaps also contributing to the transport of heat. If the scale of the flow is large, one might expect lateral inhomogeneity, but that might be extremely difficult, if not impossible, to detect, especially if the flow is direct rather than oscillatory (e.g. Gough, 2000).

One must also question whether the superficial impression one gains of the variation of the solar sound speed simply by looking at Figure 2 is really correct. There is a danger that one might imagine the true c^2 to be a smooth curve

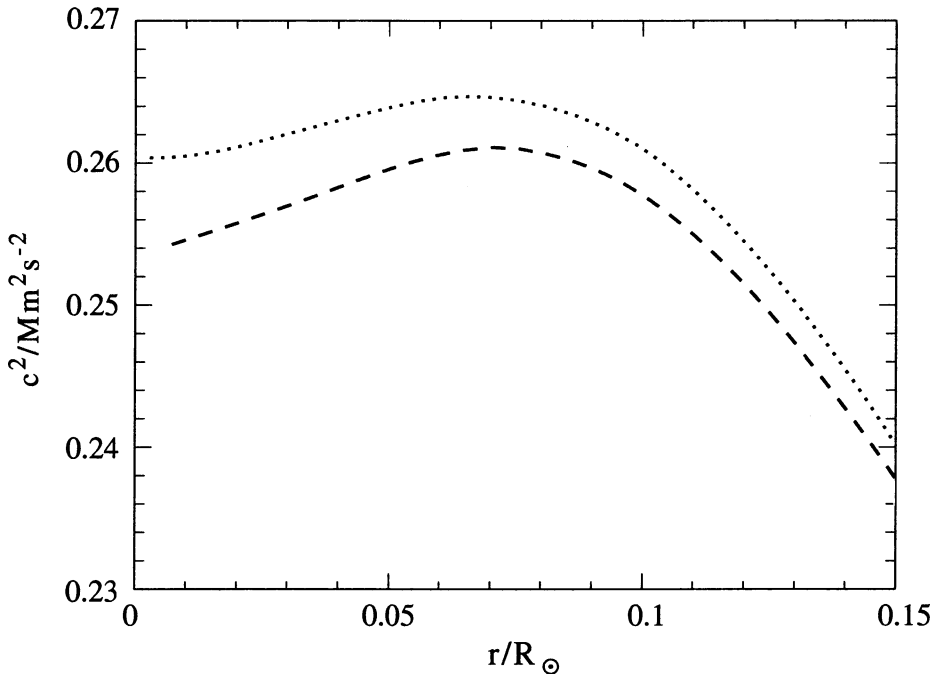


Figure 2. Square of the sound speed in the inner regions of the sun. The hashed region represents inversions by T. Sekii of frequency data from SOI/MDI and VIRGO; the dotted and dashed curves are theoretical models of J. Christensen-Dalsgaard and H. Saio respectively. The units of c are Mm s^{-1} .

drawn through the centre of the shaded region. Although the standard errors associated with the inference are rather smaller than the difference between the inference and the models, it is extremely important to remember that those errors are highly correlated, even were the errors in the raw data to have been uncorrelated (which is no doubt not the case); therefore one cannot reliably compare values at nearby points in order to estimate derivatives without having information about the error correlations (Gough, Sekii and Stark, 1996; Howe and Thompson, 1996). The effect of such correlations is to spread the influence of data errors in space, thereby obliterating some small-scale features and possibly creating others. Is, for example, the local maximum in $\delta c^2/c^2$ at $r/R \simeq 0.3$ in the left-hand panel of Fig 1 real, or is it an artefact of error correlation?

More dangerous is unaccounted correlation between the errors in the frequency data. In the absence of quantitative information, data inverters usually ignore this issue completely, and assume the frequency errors to be uncorrelated. At present they can do little else, aside from making more realistic (and necessarily more pessimistic) assessments of the likely errors in their inferences. As has been demonstrated with artificial data (Gough, 1996), the error estimates might need to be increased by a substantial factor. To be sure, in the artificial case that has been studied the correlations were deliberately somewhat contrived, in order to emphasize the point, but sometimes nature behaves sim-

ilarly; a possible real example is an apparently spurious feature in rotational splitting data encountered by Schou (1992), believed to be spurious because no other data exhibit it, yet which remains unexplained as a plausible systematic error. An essential step forward that must be given high priority, therefore, is to take due account of error correlations in the mode frequency determinations. Needless to say, effort must also be expended on the eradication of systematic errors. Without that painstaking work, the true structure and dynamics of the solar core is likely to continue to elude us.

3. Gravity modes

It is commonly, although not universally believed that the addition of only a few g modes to the helioseismic data set would substantially enhance inferences about the core. There have been reports of g-mode detections in GOLF data (Gabriel *et al.*, 1998), but these have not been convincing enough for the community to believe their implications. A more conservative report by the Phoebus group (Appourchaux *et al.*, 2000) sets an upper limit of about 1 cm s^{-1} on the g-mode amplitudes, which, disappointingly, is a factor of at least ten greater than the theoretical amplitude estimates. Perhaps methods more sophisticated than the currently used direct Fourier analysis must be brought into play.

One such method is to seek g-mode modulation of p modes by studying sidelobes in the p-mode power spectrum. The idea is to superpose the spectral lines of many p modes in the hope that if those modes are modulated substantially by only a few g modes a signature will emerge from the noise. So far as I am aware, this was tried first in 1989 on spectra of low-degree modes from single-site whole-disc data acquired in Tenerife, and subsequently by Kennedy, Jefferies and Hill (1993) on sidelobes of modes of intermediate degree, also from single-site data. Perhaps it is time to try again with the superior data from SOHO or the ground-based networks. I hasten to add that this method is not suitable for finding the stochastically excited g modes that we had hoped to have detected directly; a grave g mode with surface velocity amplitude 1 mm s^{-1} , for example, would typically produce a sidelobe peak in a p-mode spectrum with only about 10^{-16} of the power of the main peak. However, one might hope to detect the signature of a self-excited g mode that is driven in the energy-generating core, if that mode were to be nonlinearly confined to the solar interior, but then only if it has grown to a substantial amplitude without having been limited by the coupling to stable g-mode pairs as Dziembowski (1983) has suggested.

If the structural changes associated with the solar cycle are sufficient to change the g-mode frequencies by more than their damping rate, which is not unlikely if there are substantial modifications to the tachocline associated with magnetic field variation, then it would be prudent to take those changes explicitly into account when searching for g modes. One of the difficulties in doing so arises from our ignorance of the frequency changes. One can treat them as an unknown function $f(t)$ which is then adjusted to boost lines in the oscillation power spectrum in the expected frequency range. However, due caution must be exercised, because one can probably succeed in such an endeavour with a spectrum of pure noise. It is fine to use techniques such as this to improve the frequency measurements of modes that have already been identified, but to use

them as a tool for searching can be dangerous. To be even half convincing, several candidate modes would need to be found which behave in unison in a manner that makes theoretical sense. Such a method should certainly be attempted in the near future, because it may be our only hope.

4. Grave low-degree acoustic modes

We must face up to the possibility that interior g modes may not be detected and measured in the foreseeable future, and perhaps not ever. Therefore in order to refine our knowledge of the solar core we might have to rely on the grave low-degree p modes, which are the only other currently potentially diagnostically useful modes that penetrate deeply enough. One of the disadvantages of these modes is that they are influenced strongly by the spatial and temporal variations of the outer layers of the sun, which cause phase wandering of the oscillations, thereby degrading our ability to measure the frequencies. The situation could be dramatically improved by monitoring the acoustic properties of the entire surface of the sun, which requires observations not only of the near side, which we can see, but also the far side. The currently most promising approach is by telechronoseismological analysis; Lindsey and Braun (2000), motivated by a different goal, have recently succeeded in detecting distortions to the acoustic cavity brought about by far-side magnetic activity, thereby demonstrating that the necessary continuous monitoring is at least in principle possible. I shall return to this endeavour later.

5. Eigenfunction ‘distortion’

Because the sun is not spherically symmetrical, the oscillation eigenfunctions are not strictly separable in spherical polar coordinates with spherically harmonic angular dependence, as they are assumed to be in most modal analyses of helioseismic data. The difference is often thought of as a distortion (from the idealized view of the theorist). It has been calculated for theoretical models of not only the sun but also of other aspherical stars, particular roAp stars, usually as an integral part of the procedure to evaluate the frequency perturbation by rotation or a large-scale magnetic field. Woodard (1989) was one of the first to emphasize the importance to helioseismic data analysis, pointing out that the greatest deviation from the idealized harmonic structure is produced by the latitudinally varying component of the sun’s angular velocity. It is straightforward to estimate that deviation, which can be described qualitatively as a contraction at high latitudes and a slight expansion near the equator of the latitudinal scale of variation of the prograde modes, and a corresponding expansion and contraction of the retrograde modes. The deviation is not negligible for modes with degrees greater than several hundred, and for the cleanest spatial projections of the data, to which we should certainly aim, it should be taken into account even for modes of lower degree. Other deviations are produced by magnetic fields; these are less well understood, but appear to be manifest mainly as an amplitude diminution in active regions. Imperfect filtering augments sidelobes in the temporal power spectrum; the imperfections from ignoring eigenfunction distortion may not yet be as great as some other (instrumental) factors, but

it must not be forgotten in the future as the necessary improvements to the filtering are introduced.

It is perhaps worth pointing out that the other techniques for analysing solar oscillations are designed to measure the consequences of wave distortion. Indeed, the overtly local procedures, ring analysis (Patrón *et al.*, 1995) and phase analysis (Gough, Merryfield and Toomre, 1998), which concentrate on limited localized spatial regions, do so explicitly; they measure a portion of the wave field and determine a relation between frequency and spatial phase variation, with the object of determining the spatial (horizontal and vertical) variation of the properties of the underlying background state of the sun. Ring analysis has a spatial resolution comparable with the linear size of the region of averaging, typically 15° which corresponds to a distance of about $0.25R_\odot$. It is normal to assume that the mean wave properties observed in a $15^\circ \times 15^\circ$ window represent the mean conditions vertically beneath the window, which is a good approximation provided the depth at which conditions are inferred is not too great. In reality, the influence of a localized patch at depth d is spread over a surface patch of extent L , where L is roughly the distance between successive reflection points near the surface, which is approximately πd . Thus, the contribution to the horizontal blurring scale from the assumption is roughly three times the depth of the inference. Although this is smaller than the averaging scale in many of the ring analyses that have been published, it is not smaller than the resolution that some of us, perhaps overoptimistically, hope to achieve in the future from phase analysis, which is aimed at measuring the wave distortion in detail rather than merely its spatial mean; the horizontal transfer of the subphotospheric signal by the waves must therefore be accounted for in future inversions.

6. Further remarks on ring analysis

Not only does the failure to take wave distortion into account degrade the projections that attempt to isolate global modes, but it also throws away useful information that could otherwise be used in the quest to identify aspherical properties of the sun through those distortions. A case in point is the diagnosis of magnetic field. For example, for any axisymmetrical magnetic-field configuration there exists, at least asymptotically at high p-mode frequency and almost certainly for the entire acoustic frequency range, a spectrally equivalent axisymmetrical sound-speed variation — that is to say, there is always a distribution of sound-speed that yields precisely the same frequency spectrum of p modes as does a given magnetic field (Zweibel and Gough, 1995). Consequently one cannot tell the difference from frequency spectra alone, although one of the two putative causes might be more plausible than the other, although that inevitably depends on the prejudices of the beholder. It is also possible that reality is a combination of the two, together with perhaps an admixture of vertical and latitudinal advection by large-scale flow, and perturbations from smaller-scale time-dependent turbulent convective motion. Measuring wave distortion can help to reduce the ambiguity. For example, the influence of a magnetic field is locally anisotropic, leading to a direction-dependent perturbation to the wave propagation, whereas a sound-speed variation is not. Of course, there may be an appropriate anisotropic distribution of sound speed that could mimic the effect

of a magnetic field, but with adequate detail and precision of the seismic data ambiguity must surely be reduced. A thorough study of the ambiguity has not yet been undertaken.

In studying the shapes of acoustic rings — namely the power at a given temporal frequency ω as a function of horizontal wavenumber \mathbf{k} of the modes of given order n , or, equivalently, the power at constant $k = |\mathbf{k}|$ as a function of ω and direction $k^{-1}\mathbf{k}$ — it seems natural to decompose the radius $k(\psi, \omega)$ of the ring (in the first representation, or the frequency $\omega(\psi, k)$ in the second) into the Fourier components:

$$k(\omega, \psi) = k_0(\omega) \left[1 + \sum_{j>0} \alpha_j \cos(j\psi - \delta_j) \right], \quad (1)$$

where ψ is a polar angle in the horizontal coordinate system employed and k_0, α_j, δ_j are ω -dependent parameters that are to be determined from the distribution of power in the ring. To date, attention has been concentrated on α_1 and δ_1 , which characterize a displacement of the ring due to advection with velocity $(dk_0/d\omega)^{-1}\alpha_1$ of the wave in the direction $\psi = \delta_1$ by horizontal flow in the sun. The dominant distortion one might expect from the horizontal component of a magnetic field is the $j = 2$ term; in that case the direction of the field is parallel (or antiparallel) to $\psi = \frac{1}{2}\delta_2$ if $\alpha_2 > 0$. The vertical component of the field simply changes $k_0(\omega)$, as does the vertical component of velocity and a variation in sound speed, all of which are likely to be encountered in combination beneath magnetically active regions. Higher-order terms are likely to be difficult to measure from a standard ring, but the greater directional resolution with which they are associated might in the long run be achievable from phase analysis.

7. Wave interference

Real modes are not the purely sinusoidally temporally varying oscillations of simple linear perturbation theory about a smooth static background state. The constituent waves are damped and excited, and their propagation is modulated, predominantly by the turbulent convection and solar activity. Consequently their phases wander, and one cannot isolate the contributions to the seismic signal from individual modes. Moreover, by being able to observe only part of the sun's surface at each moment, one becomes subject to spatial confusion. How can the data be analysed to extract the most useful information from them? Here I shall remark briefly on the most evident problem with which wave interference confronts us, and postpone until the next section the problems of overcoming temporal variation of the waves.

If one could observe the entire surface of the sun one would be in a somewhat better position to analyse the wave field. In the absence of random excitation and damping processes, and if the background state through which the waves propagate were independent of time, one would probably not be so handicapped by an inability to observe continuously the entire solar surface; observations of only a portion of surface over a sufficiently long time interval would in principle be adequate to solve the forward problem of determining the wave field throughout the solar interior, and thereby provide the basis for an inverse analysis. (I am assuming unlimited accuracy of the observations in this idealized argument.)

It is partly because the background state varies with time in an unknown manner, and partly because the waves are excited and damped, that the forward problem is indeterminate. In practice we must, at least in a first approximation, assume the background state of the sun to be stationary when we analyse different epochs simultaneously — and even then we can hope to estimate the wave field of the forward problem only in regions from which all the waves that are emanated eventually reach the observable surface. The inevitable dissipation degrades the estimate, and since dissipation tends to increase with decreasing characteristic lengthscale of the wave disturbance, and with the distance the wave has propagated, any uncertainty in the disturbance in the observable surface, however small, leads to an uncertainty in the spatial properties of the wave field in any remote region which increases both with decreasing lengthscale of those properties and with increasing seismic remoteness of the region. Therefore the resolving power of the waves in a corresponding inverse analysis diminishes as the seismic remoteness of the region (by which I mean a measure of the distance the seismic waves must travel from the domain of observation to the region) increases. Consequently we are forced, particularly when seeking high-resolution inferences, to concentrate attention on seismically propinquant regions.

To date, all inverse analyses in which the background state is assumed to be independent of time are based on the separation of the seismic disturbance into waves or modes, namely components with time dependence of the form $e^{i\omega t}$, where ω is a (usually real) constant frequency which may be different for each component. One can thus attempt to measure the observable consequences of each component via temporal Fourier transformation of the observations. A problem that one encounters in so doing is that one cannot actually isolate every component, partly because a necessarily finite duration of observation does not permit the isolation of a single frequency ω , and partly because there are typically many wave-like components with the same frequency. One can illustrate this by considering the simple case of a group of waves with frequencies in some small interval $\Delta\omega$ propagating over a plane surface that is almost uniform and isotropic. In a simple (asymptotic) wave description one can characterize waves of frequency ω by a dispersion relation

$$\omega = W(k_i, S_{ij}) \quad (2)$$

relating the component k_i of the local (spatially varying) wavenumber $\mathbf{k}(\mathbf{x})$ at position \mathbf{x} to a seismic property $S_{ij}(\mathbf{x})$ of the background state. If the background state were perfectly uniform and isotropic, S_{ij} would be $S_0\delta_{ij}$, where S_0 is a constant and δ_{ij} is the Kronecker delta, and therefore equation (2) implies that the ray paths are rectilinear and that $\mathbf{k} = \mathbf{k}_0$ with $|\mathbf{k}_0| = k_0 = \text{constant}$, whatever the direction of \mathbf{k}_0 , with $\omega = W(\mathbf{k}_0, S_0\delta_{ij}) =: W_0(k_0, S_0)$. A small inhomogeneity and anisotropy $\Delta S_{ij}(\mathbf{x})$ (which represents, in this simple example, an asphericity in the background state of the sun which we wish to investigate) leads to a directionally dependent spatially varying wavenumber perturbation Δk_i which satisfies, to first order in the perturbations,

$$\frac{\partial W}{\partial k_i} \Delta k_i \simeq - \frac{\partial W}{\partial S_{ij}} \Delta S_{ij} \quad (3)$$

There are associated amplitude variations too. By measuring $\Delta k_i(n_j)$ associated with all directions $n_j = k_0^{-1}k_{0j}$ of the zero-order wave, one can in principle solve equation (3) for ΔS_{ij} . However, a problem arises because one cannot filter the data to isolate each of the zero-order waves, partly because the wave field also contains the 'perturbed' waves, but mainly because the observations have finite duration and are over a limited portion of the surface of the sun. Consequently each filtration of the signal contains a packet of waves, which interfere to cause phase and amplitude variations in addition to those produced by ΔS_{ij} . Can they be distinguished from the ΔS_{ij} -induced variations? One might suspect that they can. The interference patterns propagate across the field of view, and if $\Delta\omega$ is small they do so with an effective group velocity v_i , whereas the ΔS_{ij} -induced variations are stationary. So perhaps they can be eliminated.

Most attempts to achieve that elimination are based simply on averaging the data over a long interval of time, for the interference pattern eventually averages to zero whereas the pattern produced by the variation ΔS_{ij} of the background state (which is presumed to be static) survives. Evidently the averaging interval must be very much greater than the time it takes to cross the domain of observation at the effective group velocity. It is therefore important to assess how long that time is. We know also that in reality the background state is changing with time, so one cannot afford to average over more than the characteristic timescale of that variation.

If all the waves were propagating in the same direction n_i , then the effective group velocity would be just what we normally regard as group velocity, namely $v_i = (\partial W_0/\partial k_0)n_i$. For the surface pattern of high-degree acoustic waves in the sun, for example, its magnitude would be approximately half the horizontal phase speed of the zero-order wave (namely, about $\frac{1}{2}\omega k_0^{-1}$). However, because one cannot successfully filter out waves in a unique direction, one is left with a fan about some direction \hat{n} , with apparent wavenumbers $\hat{n}\cdot k$ in that direction. Thus, there is a packet of waves with a nonzero range of apparent wavenumber in some direction, and, because of the near isotropy of the background state, all with essentially the same frequency. Thus their effective group velocity is zero, and therefore their interference pattern cannot be distinguished from the pattern produced by ΔS_{ij} ; accordingly one might expect them to survive the temporal averaging typically carried out in current data analyses.

Fortunately all is not necessarily lost. Neighbouring fans of rays separated in angle by the angular extent or more of both fans produce independent interference patterns, so provided the quality and extent of the data permit separation into narrow fans (although the data must permit this, the analysis of them need not explicitly do it) the interference pattern varies predominantly on a small (angular) scale. Therefore appropriate spatial smoothing should filter it out. It would also filter out any small-scale variation of ΔS_{ij} , of course, but at least one has the hope of accessing the large-scale variation. Learning how to realize that hope is an important task for the future.

8. Temporal variation

I shall divide the temporal variation of the background state into two components, the first with characteristic timescales less than or comparable with the

frequencies of the seismic waves, the second with timescales much longer. The former produces what we normally consider to be the excitation and damping of the seismic modes, the latter produces long-term modulation of the amplitude and phase, and we can regard it as a fluctuation principally of the propagation characteristics of the acoustic cavity (I shall not discuss gravity modes here). It is generally believed — and indeed there is considerable evidence for the belief — that the seismically most important fluctuations take place in the immediately subphotospheric layers of the sun: the rapid fluctuations are the turbulence in the upper superadiabatic boundary layer of the convection zone, which acts principally as a source of excitation, absorption and scattering of acoustic waves, the slow fluctuations are associated mainly with the so-called magnetic activity.

Because seismic modes experience these fluctuations in what appears on a global scale to be a thin (essentially two-dimensional) spherical shell, all horizontally similar modes (modes of the same degree and azimuthal order) are 'forced' by the same perturbation from the stationary state, to within a calculable constant factor which depends on the radial order of the mode. One might therefore hope to combine the data describing a group of horizontally similar modes to extract the influence of the temporal variation. The usefulness of so doing would be two-fold: firstly, account could be taken of that influence on the acoustic modes to infer how those modes would have behaved in the absence of the forcing, which simplifies the task of subsequent inversion (indeed, simplifies it to that which is performed at present — namely, to invert under the assumption that the background state is steady), and secondly, knowledge of the forcing provides some information about the temporal variations of the background state which could be used to improve our knowledge of the properties of the turbulence and the magnetic activity.

It is unlikely that in the immediate future any substantial coherent influence of the rapid fluctuations on the amplitude and phase wandering of the modes can be effectively removed to infer how the modes would have behaved in the absence of those fluctuations. That is a pity, because had it been possible it would have provided the means to make substantial improvements on our estimates of the free-oscillation frequencies. That would have been particularly important for low-degree acoustic modes, which in the absence of gravity-mode observations are needed for diagnosing the structure of the sun's energy-generating core.

Part of the reason for the elimination not to be imminently possible is that individual modes in a horizontally similar group cannot currently be observed separately: the single observed signal is from the sum of all the modes of the group. More important, perhaps, is that there are only a few modes in the group, yet there are many degrees of freedom in the random forcing function, about as many per period of oscillation as the number of granules exciting the modes: millions. The wandering of the amplitude and phase, which, because the so-called lifetimes of the modes are long compared with their oscillation periods, can be thought of as being slow; it is the product of a temporal average of the forcing, weighted by an appropriate measure of the instantaneous perturbation characterizing the oscillation. Since the modes of the group oscillate approximately sinusoidally in time, with frequencies which, at low degree, are roughly in harmonic sequence, each average of the forcing can be thought of crudely as a Fourier component, and is independent of any other. Therefore the forcing

of one mode provides essentially no information about the forcing of another; there is no shared information that can be extracted from the oscillation data. At present, therefore, all that can be done is to do as we are doing: temporally average the data in some manner (typically by analysing either a single Fourier power spectrum of the entire data set or an average of the Fourier spectra of subsets) to eliminate the gross effects of the fluctuations, a process which reduces the errors in the frequency determination only as the inverse square root of the duration of the observations.

One can conceive, perhaps, of the possibility in the future of making more detailed observations of the oscillations in the photosphere and the overlying atmosphere to use the different properties of the modes (height variation, amplitude and phase relations between, say, displacement and temperature and density perturbations, etc.) to distinguish their individual contributions to the signals. That is bound to offer some improvement, but it cannot overcome the fundamental difficulty of the nearly orthogonal sampling of the forcing. Only if the forcing can be inferred independently (through observations of the granular motion, for example, over the entire surface of the sun, coupled with a reliable theory of how that relates to the forcing) might we be able to make really substantially improvements in our determination of mode frequencies.

There is more hope of accounting for the slow fluctuations. Experiments with artificial data have shown that if the signal from a horizontally similar group of modes is good enough for the contributions from each individual mode to be detectable above both the noise and the contributions from the other modes of the group, then, provided the relative strengths of the influence of the fluctuations on the mode properties is known, it is possible both to determine what those fluctuations are and to obtain estimates of the mean mode frequencies which are significantly more accurate than corresponding estimates obtained by mere temporal averaging (Chang, Gough and Sekii, 1995). This can be achieved only for modes whose 'lifetimes' (i.e. characteristic times over which phase coherence, corrected for the slow fluctuations, is maintained against the rapid stochastic forcing due to the turbulence) are greater than the characteristic timescale of variation of the slow fluctuations, which is the case only for low-degree g modes and grave low-degree p modes. However, these are the very modes for which high accuracy is required for determining the structure of the sun's core.

Equally interesting is to use a proxy of the slow fluctuations in the properties of the seismic cavity to obtain an independent estimate of the phase and amplitude wandering of the grave low-degree modes, and to incorporate that information into the procedures for analysing the seismic data for the purpose of discovering new modes. Thus, rather than merely computing Fourier spectra, one attempts to project models of the fluctuating modes onto the data. Should this be feasible, it might then be possible to make the signals from new low-frequency modes emerge out of the solar background noise. Woodard *et al.* (1991), Bachmann and Brown (1993) and Jimenez-Reyes *et al.* (1998) have shown how frequency variation is correlated with magnetic activity; the magnitude of the variation is reported to be independent of the timescale of variation of magnetic activity, and it appears to be possible to model the variation with an appropriate activity-generated phase shift imposed on acoustic waves in the up-

per reflecting layers of the acoustic cavity. This is borne out by Duvall's (1995) measurements of the influence of sunspots on acoustic phase travel time, using time-distance techniques, which is more-or-less consistent with the measured perturbation to the frequencies of global modes (Cunha, Brügger and Gough, 1998). Therefore an observed measure of magnetic activity can be used as the acoustic proxy. Evidently, only variations on timescales greater than about two weeks can be taken into account, because it will be necessary to interpolate the measurements of activity on the visible disc of the sun to estimate the activity on the back side. This may not be adequate for discovering new modes.

An alternative and expectedly superior procedure is to measure the acoustic properties of the cavity directly, by monitoring the phase shifts produced over the entire surface of the cavity using time-distance techniques. As I have already pointed out, Lindsey and Braun (2000) have provided evidence that it is feasible. It must be realized, however, that its utility in the quest for discovering new lower-order modes is predicated on the presumption that the principal agents that modulate the phases of the low-frequency modes are the same as those that principally modulate the higher-frequency modes. If not, it does not follow that the time-distance data cannot be used for the purpose, but it will evidently be much more difficult to use them.

9. Error Correlation

I return to this subject because it is so important. Current techniques for analysing Fourier power spectra of helioseismic data to determine the oscillation frequencies of modes yield results with errors that are not independent of one another. Estimates of the magnitudes of the random errors are provided by the data analysts, but the error-correlation matrix describing their interdependence almost invariably is not. As I pointed out earlier, for want of an alternative, inverters are forced to resort to assuming the frequency errors to be independent. That can lead to erroneous results, with which are associated unrealistically small error estimates that can falsely appear to indicate that spurious features in the inversions are extremely unlikely to have been the outcome of random data errors. It is therefore extremely important that more work be carried out on investigating what the error correlations might be. The work is laborious, and it may at first sight seem to be more important to use one's limited resources to obtain new results rather than to use them to make the expensive estimates of the error properties of the results one has just obtained. That is no doubt correct in the early phases of an investigation. But in more advanced phases, such as the one helioseismology is entering in the new millennium, very much more subtle features of the data than have been sought in the past are being investigated to discover fine details of the structure and motion in the solar interior, and these are much more likely to be severely contaminated by correlated errors. The inversion procedures have no internal consistency checks for some aspects of the correlations, other than to compare inversions of totally independent data sets. Therefore it really is of paramount importance to do one's (reasonable) utmost to estimate the correlations from the procedures that generate them.

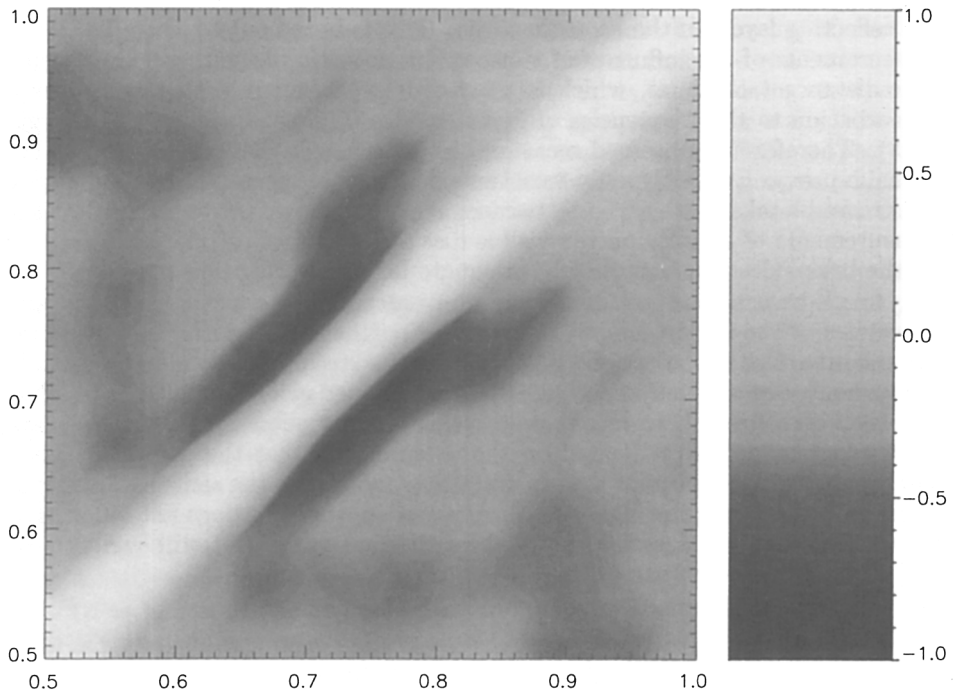


Figure 3. Correlation between optimally localized averages (of Gough, Sekii and Stark, 1996) at different radii of the spherically averaged angular velocity in the outer half (by radius) of the sun, computed under the assumption that the frequency data errors are independent of one another.

Of course, additional correlations are introduced by the inversion procedures. These are easily computed from linearized procedures (which most are), and have been discussed a little in the literature. The typical structure of a correlation matrix for, say, averages in radius of the spherically symmetrical component of angular velocity inferred from p-mode frequency splitting with uncorrelated errors is illustrated in Figure 3. It is dominated by a positive diagonal strip, representing correlations between inferences at neighbouring radii, and two flanking negative strips. There are also strips near the edges close to the deepest inversion radius, representing contamination of the deepest values by errors at all depths. But when the frequency data errors are correlated, as in practice they are, the matrix could look rather different.

10. Telechronoseismology

Although helioseismological time-distance analysis is in a very primitive state compared with global-mode analysis, it is advancing rapidly, and is becoming a very powerful new tool. It is particularly well suited to measuring horizontal inhomogeneity, because it combines temporal information (phase and group propagation times) with information about spatial distortion. Without the spatial information there are many aspects of the geometry of the inhomogene-

ity that are seismically inaccessible. The most obvious aspect is simply global orientation: a vibrating object oscillates at the same frequencies whatever its orientation relative to the beholder. It is harder to make further simple precise statements, except to point out that distinct isospectral configurations are known to exist for even the simplest of configurations in more than one dimension (e.g. Driscoll, 1997), as was discovered in answering the famous question posed by Mark Kac in a lecture in 1966: 'Can one hear the shape of a drum?'

Most time-distance analysis has been accomplished in terms of ray theory in its very simplest form. Group and phase travel times for propagation between different points on the solar surface are related to the structure and motion of the sun's interior via an asymptotic dispersion relation describing propagation along a ray, the influence of background velocity, whether it be large-scale rotation or meridional flow or smaller-scale convection, being regarded simply as straightforward advection. As in the case of global-mode analysis, motion and asphericity in structure of the background state can be investigated by perturbation theory. When it is phase times that are measured, the analysis is rendered relatively simple by there being no necessity to perturb the ray paths, because, in the light of a generalization of Fermat's principle, small perturbations to the path of integration from the true ray path lead to only second-order perturbations to the propagation time. There is no analogous variational principle for group propagation (notwithstanding claims to the contrary in the astrophysical literature), so formally integration should be carried out along the true, perturbed ray paths. Be that as it may, the group measurements are possibly insufficiently accurate for such a detail to matter at the moment, although a careful investigation of the issue is yet to be undertaken.

The most spectacular of the first telechronological results are representations of sound-speed fluctuations and fluid flow in the convection zone, both in the quiet sun (Kosovichev and Duvall, 1997) and in the vicinity of sunspots (e.g. Kosovichev, 1996; Kosovichev, Duvall and Scherrer, 2000). They were obtained by regularized least-squares data fitting to travel times along (unperturbed) rays, using a representation of the fluctuations by piecewise constant functions on a three-dimensional grid. Although the procedure was tested with artificial data, and the results are certainly plausible, the conclusions have been challenged. They will surely be checked in the near future, although the computational task is currently heavy for most institutions.

A computationally less demanding procedure is to construct a virtual acoustic lens, and concentrate on limited regions at a time. This has been called acoustic imaging. Lindsey and Braun (2000), Chou, Sun and Chang (2000) and Chou and Duvall (2000), amongst others, have adopted this approach, with promising results. Both group propagation and phase propagation have been considered, and the relation between them has been discussed in idealized terms.

One of the issues that has not been investigated with care is how to take due account of the fact that the observations are made in the evanescent zone (of the dominant constituent waves) above the upper caustic envelope of the rays. If the sun were oscillating as a superposition of (undamped and unexcited) normal modes of a stationary background state, which it is not, one might imagine that the visible photospheric regions oscillate precisely in phase with the propagating waves at the caustic surface which bounds the region of propagation, and indeed

that would be the case if a single discrete frequency could be isolated from the rest of the signal. But in practice one must contend with wave interference within a group of waves, and take account of the time it takes for the information to propagate through the evanescent region.¹ This is an important task for the future.

Only a little work has been done on studying the influence of the acoustic disturbance off the principal ray path. In reality, the relation between the disturbances at two points A and B on the sun's surface depends on all acoustically accessible points in the acoustic cavity (and beyond), and not just those joined by the principal ray paths, although the predominant influence often comes from the region occupied by a tube of rays in the neighbourhood of the principal ray (cf. Stark and Nikolayev, 1993). Figure 4 illustrates such a tube (with three remote 'reflections' near the surface) pertinent to investigating active regions on the back side of the sun. Rather than reacting to conditions just at isolated points on the surface (assuming, for simplicity, that there are no substantial unknown perturbations in the interior) the rays in the tube sample a finite patch, predominantly the inner Fresnel zone, and therefore the reaction to a small sunspot S which one might be seeking is diluted in inverse proportion to the size of the patch. The size of the patch decreases as the number of surface 'reflections' of the tube between A and B increases, which is associated with an increase in sensitivity to the spot; but so too does the degradation of the signal due to absorption, scattering and emission by the turbulence of the superadiabatic boundary layer of the convection zone. The optimal tradeoff appears to be about three or four reflections (cf. Lindsey and Braun, 2000), although once again it is the case that we have insufficient experience to know the truly optimal procedure. It is not unlikely that combining signals from several different tubes of rays would improve matters further.

¹It has often been remarked to me that waves do not propagate through evanescent regions with a finite speed, and that information is received instantaneously throughout the region above the upper (reflecting) caustic surface that bounds the acoustic cavity in the sun. There is a sense in which the first part of the statement is correct; but the second part is wrong. When a packet of (linear) waves arrives at the (apparent) caustic, the disturbance certainly does propagate upwards, changing its shape on the way, the front propagating with the local speed of sound; acoustic information cannot propagate faster. Thus there is definitely a time delay between the arrival of a disturbance at the caustic and its reception in the visible region above. The phase, on the other hand, is either advanced or retarded, according to whether the amplitude of the interference pattern is rising or falling. One way of looking at the situation is to imagine the disturbance to be regarded as a Fourier integral in time. The integrand contains components at all frequencies, and components with frequency above the acoustic cutoff can propagate through what I regarded to be the evanescent region. That is why I called the caustic 'apparent', for the region above is not evanescent to the entire disturbance. The propagation speed is a function of frequency, which is why the disturbance changes its shape (as indeed it does also in the region of propagation, particularly in the vicinity of the upper caustic where the acoustic cutoff frequency is comparable with the central frequency of the packet). The invariance with height in the evanescent region of the temporal phase of an oscillation in a pure undamped mode does not imply that phase information is transmitted instantaneously; indeed, for the oscillation to be of a genuinely unique frequency it must have existed for all time, and the motion in the evanescent region can formally be considered to be the response to the entire history. When the oscillation is not strictly periodic, information propagating upwards can interfere to cause phase changes in either direction, and is such that the long-term average of the change vanishes for statistically steady amplitude modulation.

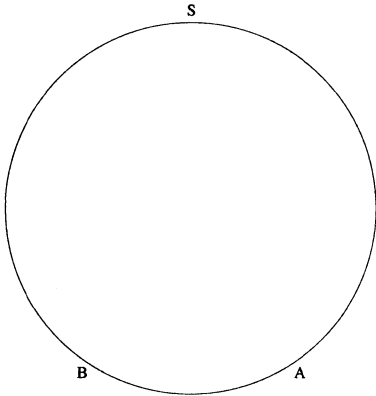


Figure 4. Projection onto a poloidal plane of a tube of rays acoustically connecting two points A and B on the surface of the sun. The principal information about the reflecting properties of the surface near S (and near the other reflecting regions) is carried in this tube. The signal containing information about the properties in the inner Fresnel zone of the point S is enhanced relative to other points on the surface by averaging over the set of tubes obtained by rotating the figure about an axis passing through the centre of the sun and through S.

The most ambitious programme would be to attempt an inverse problem based on the complete solution to the three-dimensional wave equation with a temporally varying inhomogeneous background in terms of the disturbance on the observable surface. Computationally this is at present a formidable undertaking, but it will not be so in the distant future. I suspect, however, that when it becomes possible we shall consider ourselves to have entered yet another era of helioseismology.

Acknowledgments. I am grateful to D. Bayston, G. Houdek and R. Sword for their assistance with the production of the paper.

References

- Appourchaux, T., Fröhlich, C. Anderson, B., Berthomieu, G., Chaplin, W.J., Elsworth, Y., Finsterle, W., Gough, D.O., Hoeksema, J.T., Isaak, G.R., Kosovichev, A.G., Provost, J., Scherrer, P.H., Sekii, T. and Toutain, T., 2000, *Astrophys.J.*, 538, 401–414
- Bachmann, K.T. and Brown, T.M., 1993, *Astrophys.J.*, 411, L45-48
- Chang, H-Y., Gough, D.O. and Sekii, T., 1995, *Proc. Fourth SOHO Workshop: Helioseismology*, ed. J.T. Hoeksema, V. Domingo, B. Fleck and B. Battrick, ESA SP-376, vol 2, 175–178
- Chou, D-Y. and Duvall, T.L.Jr, 2000, *Astrophys.J.*, 533, 568-573
- Chou, D-Y., Sun, M-T. and Chang, H-K., 2000, *Astrophys.J.*, 532, 622-627
- Christensen-Dalsgaard, J. et al., 1996, *Science*, 272, 1286-1292
- Cunha, M.S., Brüggen, M. and Gough, D.O., 1998, *Proc. SOHO6/GONG98 Workshop: Structure and dynamics of the interior of the sun and sun-like stars*, ed. S. Korzennik and A. Wilson, ESA SP-418, 905-910
- Driscoll, T.A., 1997, *SIAM Rev.*, 39, 1-17
- Duvall, T.L. Jr, 1995, *GONG '94: Helio- and astero-seismology from the earth and space*, ed. R.K. Ulrich, E.J. Rhodes and W. Däppen, *Astron. Soc. Pacific Conf. Ser.*, 76, 465-474

- Dziembowski, W.A., 1983, *Solar Phys.*, 82, 259-266
- Gabriel, A.H., Turck-Chieze, S., Garcia, R.A., Palle, P.L., Boumier, P., Thiery, S., Grec, G., Ulrich, R.K., Bertello, L., Roca Cortes, T., Robillot, J.-M., 1998, SOHO6/GONG98, Workshop: Structure and Dynamics of the Interior of the Sun and Sun-like Stars, ed. S. Korzennik and A. Wilson, ESA SP-418, 61-66
- Gough, D.O., 1996, The structure of the Sun, ed. T. Roca Cortés and F. Sánchez, Cambridge University Press, Cambridge, 141-228
- Gough, D.O., 2000, *Solar Phys.*, in press
- Gough, D.O., Merryfield, W.J. and Toomre, J., 1998, *Astrophys. J.*, 501, 882-894
- Gough, D.O., Sekii, T. and Stark, P.B., 1996, *Astrophys. J.*, 459, 779-791
- Howe, R. and Thomson, M.J., 1996, *MNRAS*, 281, 1385-1392
- Jiménez-Reyes, S.J., Regulo, C., Pallé, P. and Roca Cortés, T., 1998, *Astron. Astrophys.*, 329, 1119-1124
- Jordinson, C. and Gough, D.O., 2001, The Impact of Large-Scale Surveys on Pulsating-Star Research, ed. L. Szabados & D.W. Kurtz, *Astron. Soc. Pacific Conf. Ser.*, in press
- Kajita, T., 1999, *Nucl.Phys.*, B (Proc.Suppl), 77, 123
- Kennedy, J.R., Jefferies, S.M. and Hill, F., 1993, GONG 1992: Seismic investigation of the sun and stars, ed. T.M. Brown, *ASP Conf. Ser.*, 42, 273-276
- Kosovichev, A.G., 1996, *Astrophys.J.*, 461, L55-57
- Kosovichev, A.G. and Duvall, T.L. Jr, 1997, SCORE'96: Solar convection and oscillations and their relationship, ed. F.P. Pijpers, J. Christensen-Dalsgaard and C.S. Rosenthal, Kluwer, Dordrecht, p. 241-260
- Kosovichev, A.G. et al., 1997a, Sounding solar and stellar interiors, ed. J. Provost and F-X. Schmider, Kluwer, Dordrecht, *Proc. IAU Symp.* 181, 203-210
- Kosovichev, A.G., Duvall, T.L.Jr and Scherrer, P.H., 2000, *Solar Phys.*, 192, 159-176
- Lindsey, C. and Braun, D.C., 2000, *Science*, 287, 1799-1801
- Patrón, J., Hill, F., Rhodes, E.J.Jr, Korzennik, S.G. and Cacciani, A., 1995, *Astrophys. J.*, 455, 746-757
- Schou, J., 1992, Ph.D. dissertation, Århus Universitet
- Schou, J. et al., 1998, *Astrophys. J.*, 505, 390-417
- Stark, P.B. and Nikolayev, D.I., 1993, *J. Geophys. Res.*, 98, 8095-8106
- Woodard, M.F. Libbrecht, K.G., Kuhn, J.R. and Murray, N., 1991, *Astrophys. J.*, 373, L81-84
- Zweibel, E.G. and Gough, D.O., 1995, Proc. Fourth SOHO Workshop: Helioseismology, ed. J.T. Hoeksema, V. Domingo, B. Fleck and B. Battrick, ESA SP-376, vol 2, 73-76