A REVIEW OF THE THEORY OF SOLAR OSCILLATIONS AND ITS IMPLICATIONS CONCERNING THE INTERNAL STRUCTURE OF THE SUN

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

ABSTRACT

The method by which solar models are computed and calibrated is summarized. It is explained how oscillations of high and of low degree measure two quite different aspects of those models. However, when each is compared with observation, essentially the same model is selected. That model has an initial helium abundance: Y = 0.27±0.01.

The model so selected does not agree with all observations, for it predicts a neutrino flux some three times higher than that observed. Moreover, there are disturbing systematic differences between the theoretical eigenfrequencies and observation. However, there is good reason to believe that future observations of low-order modes of low and possibly intermediate degree will explain these differences, and lead to a more accurate description of the solar interior.

Evidence for the measurement of rotational splitting in solar oscillations is briefly reviewed. At present, we are uncertain of the precise distribution of angular velocity in the solar interior. Nevertheless, the evidence seems to suggest a central rotation considerably faster than the rotation of the photosphere.

ON THE COMPUTATION OF SOLAR MODELS

Solar models are computed under the simplest assumptions of stellar evolution theory: the sun is assumed to be spherically symmetrical, and in hydrostatic equilibrium. The direct effects of rotation and magnetic fields on that equilibrium are usually ignored, and so is the balance (or imbalance, rather) of the horizontal components of the centrifugal and the Lorentz forces. The magnitudes of those forces are probably very much smaller than the magnitudes of the pressure gradient and gravity, and so their neglect in the hydrostatic equation is a good approximation. However, the large-scale meridional Eddington-Sweet circulation that ensues from the slight horizontal imbalance could be more serious. If the sun rotated more-or-less uniformly at the photospheric rate, the meridional circulation would have a timescale of
roughly $10^{12}$ years in the absence of a magnetic field. This is substantially greater than the age of the sun, and is the reason for ignoring the flow. However, this timescale is inversely proportional to centrifugal force, and hence to the square of the angular velocity $\Omega$; so if $\Omega$ were only some ten times greater in the solar interior, the circulation timescale would be comparable with the sun's age. Moreover, regions in which $\Omega$ varies rapidly could produce substantial but localized laminar circulation even if the magnitude of $\Omega$ were not much greater than the photospheric value. The effects on the circulation of the magnetic field are difficult to estimate.

The importance of the meridional circulation to stellar evolution is the material mixing it produces. In so-called standard evolution theory, mixing is accounted for only where there is convection. In solar models, convection occurs only in an outer shell about $0.3 R_\odot$ deep. Therefore the products of the nuclear reactions in the core are presumed to remain in situ. Evidently, if meridional circulation mixed the core with its surroundings, evolution would be modified. In particular, one would expect the replenishment of hydrogen in the core to prolong the Main-Sequence hydrogen-burning phase of the sun's evolution. If that were typical of other stars too, it would mean that we have underestimated the ages of the globular clusters (and of the galaxy), for those estimates are made by comparing the observed point of departure from the Main Sequence with the predictions of standard theory. A severe underestimate could have significant cosmological repercussions.

Aside from the large-scale meridional flow, material mixing could also occur as a result of small-scale, possibly turbulent motion. This could result from rotational or magnetic instabilities, or from wave breaking. The consequence of such mixing on estimates of Main-Sequence lifetimes could be similar to that of the large-scale meridional flow. However, the possibility of instabilities raises another issue, which deserves some attention.

In standard theory, the sun is in thermal balance: the rate at which energy is generated in the interior is precisely balanced by the observed rate at which energy is radiated from the surface. This balance is a consequence of the fact that the thermal adjustment time, which is about $3 \times 10^7$ y, is very much less than the age of the sun, which is usually taken to be $4.70 - 4.75 \times 10^9$ y. It is normally presumed that the sun is globally stable, so that any thermal imbalance present initially has had plenty of time to have been annulled. However, it has been suggested that thermal balance has been upset more recently, as a result of instabilities initiated by wave breaking or by shear. Resultant luminosity variations have been associated with terrestrial ice ages, and since the Quaternary Ice Age commenced probably no more than $3 \times 10^6$ y ago, it is possible that at the present time thermal balance in the sun has not yet been restored.

Another assumption of standard theory is conservation of material. Mass loss by the solar wind is currently negligible, and probably has always been so. Accretion is unlikely to have changed the sun's mass significantly. However, it has been conjectured that the gravitational energy released might
sometimes enhance the luminosity. Moreover, the material accreted could mask the chemical composition beneath the convection zone.

It is evident that even if we do not question the microphysics of the sun: the equation of state, the opacity and the nuclear reaction rates, and if we do not question the sun's age, we nevertheless have good grounds for doubting the predictions of standard solar models. There is reason to believe that many of the uncertainties will be resolved by an analysis of solar oscillations, which is the subject of my lecture. But first I should describe the standard solar models, to provide a basis for that analysis.

THE CALIBRATION OF STANDARD SOLAR MODELS

Standard solar models are presumed to have been initially chemically homogeneous. This is usually justified by the likelihood that the sun experienced a fully convective Hayashi phase before arriving on the Main Sequence. The initial composition is parameterized by the relative abundances by mass X, Y, and Z of \(^1\)H, \(^4\)He, and of all other elements combined. The abundance Z is small. Thus

\[
X + Y = 1 - Z = 1
\]

During evolution Y increases in the core, mainly at the expense of X, and there is some readjustment of the constituents of Z. In the envelope, X, Y and Z remain constant.

There is an additional parameter of the solar model, which controls the structure of the convection zone. There is no satisfactory theory of stellar convection, though this is probably not a very serious matter when it comes to estimating the gross characteristics of the sun. The reason is that apart from in a thin boundary layer at the top of the zone, convection is so efficient that the stratification is very close to being adiabatic. Thus the role of convection theory is simply to describe the upper boundary layer. Since that layer is too thin for its detailed structure to have much bearing on the structure of the rest of the star, the only property of importance (so far as my discussion today is concerned) is the jump of temperature (or entropy) across it. That can be specified by a single parameter. In practice one adopts a local relation between heat flux and temperature gradient, usually based on a mixing-length formalism, which contains an adjustable parameter, say \(\alpha\). The solar model is insensitive to the details of the formula, provided heat flux is an increasing function of temperature gradient.

The principal calibration of the model is to reproduce the observed present luminosity \(L_\odot\) and radius \(R_\odot\). For that one has available three independent adjustable parameters: Y and Z, say, and \(\alpha\). Hence, if a solution exists, there must be a single infinity of them. Solutions do exist, and I shall label them with the helium abundance Y.
Fig. 1. Schematic representation of the temperature distribution in two solar models. Model A is the calibrated model, with $Y = 0.27$, and is represented by the continuous line. The dashed line represents Model B, with $Y \ll 0.2$. The diagram is not to scale; in particular the structure in the convective boundary layer has been exaggerated in order that it can be seen.

The next task is to select the best solution, and for this we must use another observation. One might have chosen the neutrino flux, which is essentially a measure of central temperature, but that leads to a problem which I shall mention later. To set the scene I shall first describe how the calibrated models vary as $Y$ is changed.

In Figure 1 is plotted schematically the temperature $T$ as a function of radius $r$. The temperature gradient vanishes at the center, and its magnitude increases outwards, owing mainly to a rise in opacity. Eventually the gradient becomes so steep that the criterion for convective stability is violated. Beyond that level the temperature stratification is essentially adiabatic, almost to the photosphere. The gradient is substantially superadiabatic in the thin convective boundary layer just beneath the photosphere, which is calibrated by varying $\alpha$ to yield the correct value of $T$ in the photosphere. Roughly speaking, $T$ varies linearly with depth in the adiabatically stratified region.

Consider now the effect on a given stellar model of increasing $Z$. Since $Z \ll 1$, it has only a minor influence on the equation of state. Its main influence is on the opacity. Increasing $Z$ increases the opacity. This inhibits the radiation flux beneath the convection zone, and the model must expand to compensate. A more distended solar interior has less weight, since the material is in a lower gravitational field, and it can therefore be supported at a lower temperature. Thus density and temperature are reduced as $Z$ is increased; there is a consequent reduction in the thermonuclear energy generation rate, and hence in the luminosity $L$ of the model.
The effect of increasing $X$ is similar, but for a different reason. Increasing $X$ increases the number of free electrons per unit mass, and so increases the opacity. Thus, as I have already explained, there is a tendency to reduce $L$, though not by as much as the reduction produced by a comparable increase in $Z$. What is more important is the influence on the equation of state. Higher $X$ implies more particles per unit mass, and a greater pressure at a given temperature. This also leads to an expansion, a reduction in internal temperature, and a consequent reduction in $L$.

Since $L = L_0$ is a condition that must be satisfied by any solar model, increasing $X$ (essentially at the expense of $Y$) must be compensated by an appropriate decrease in $Z$. Thus, in a sequence of calibrated solar models, models with lower $Y$ have lower $Z$ and lower opacity. They also have lower central temperatures. Notice that such models have more hydrogen, which tends to increase nuclear reaction rates, thus counteracting the decrease brought about by the reduction in the central density and temperature.

So far, I have said nothing about the effect on the photospheric radius $R$ when $X$ or $Z$ are changed. The reason is that $R$ can be adjusted by changing $\alpha$. Decreasing $\alpha$ decreases the efficacy of convection. This is similar to increasing opacity, and leads to a distention of the star. However, the efficacy of convection is important only in the thin outer superadiabatic boundary layer, and involves only a minute fraction of the solar mass (there is less than $10^{-7}$ of the sun's mass above the level at which the deviation from adiabatic stratification exceeds 1 per cent). Consequently the hydrostatic modification that is experienced by the core is hardly noticeable, and $L$ is almost unchanged. Thus the calibrations of luminosity and radius are almost independent: chemical composition is adjusted to maintain luminosity, and $\alpha$ is adjusted to maintain radius.

In Figure 1 the temperature distribution of what one might call a normal model (Model A), indicated by the continuous line, is compared with that of a model with low $Y$ and $Z$ (Model B). In Model B opacity is lower, and so therefore is the temperature gradient in the radiative interior. Thus eventually the two curves cross: the temperature in Model B is greater in the outer layers. Another consequence of the lower temperature gradient in Model B is that radiative equilibrium extends to greater radii before the convective stability criterion is violated: Model B has a shallower convection zone than Model A.

THE SOLAR NEUTRINO PROBLEM

A model from a calibrated sequence could be selected by choosing the one (there is only one) with the appropriate neutrino flux (cf. Bahcall and Davis 1976). But then a problem is encountered: the model so selected has a chemical composition which is at variance with astronomers' preconceived ideas. If the sun were typical of stars of its age, one would expect $Z \approx 0.02$ and $Y \approx 0.25$. In contrast, a model much lower in heavy elements would be
required to reproduce the observed flux (Abraham and Iben 1971, Bahcall and Ulrich 1971). Though a drastic modification to the uncertain opacity calculations might conceivably have led to a more reasonable value of \( Z \), the calibrated value of \( Y \) is robust. And that value is less than 0.2, which is lower than many estimates of the yield from the Big Bang.

Another problem was that Bahcall and Ulrich were unable to obtain a solar model with both the correct neutrino flux and a sufficiently large radius. However, that problem was removed when, at the suggestion of Joss (1974), helium-deficient models were constructed with convection zones rich in heavy elements (Christensen-Dalsgaard et al. 1979), the enrichment having been presumed to have been provided by matter accreted during Main-Sequence evolution. Those dirty models, which, like their uncontaminated counterparts, have shallower convection zones and lower central temperatures and densities than Model A, now seem to be ruled out by the oscillation data.

**SOLAR OSCILLATIONS**

To a first approximation the sun may be considered to be spherically symmetrical and, on the characteristic dynamical timescale, the basic structure is independent of time. Hence the linear eigenfunctions of oscillation are separable in spherical polar coordinates \((r, \theta, \phi)\) about the center of the sun, and in time \( t \). In particular, the displacement eigenfunction \( \xi \) may be written:

\[
\xi(r, t) = \text{Re}\left[ \xi(r) S_{\lambda m}, \eta(r) \frac{d}{d\theta} S_{\lambda m}^*, \text{im} \eta(r) \csc \theta S_{\lambda m} \right] e^{i\omega t} \]  

(2)

where

\[
S_{\lambda m}(\theta, \phi) = P_{\lambda}^{m} \cos \theta e^{im\phi} \]  

(3)

and \( P_{\lambda}^{m} \) is the associated Legendre function of the first kind. For given \( \lambda \) there is a sequence of eigensolutions \((\xi, \eta; \omega)\) which can be arranged in order of increasing \( \omega \) and labeled with an integer \( n \). The sequence is divided into two regimes, with \( \xi \) and \( \eta \) becoming more corrugated functions of \( r \) as \( \omega \) becomes very small or very large. Indeed, it is possible to choose the origin of \( n \) such that \(|n|\) is the number of zeros in \( \xi \) as \(|n| \rightarrow \infty \). Then \(|n|\) is usually called the order of the mode. [Sometimes \( n \) is used for order, to distinguish between the two types of mode.] The degree \( \lambda \) of the spherical harmonic \( S_{\lambda m} \) is called the degree of the mode, and the order \( m \) of \( S_{\lambda m} \) is the angular, or azimuthal, order of the mode.

The order \( n \) measures the vertical component of the wave number of the oscillation; \( \lambda \) measures the total horizontal wave number, and \( m \) its azimuthal component. Evidently \( m \) depends on the orientation of the axis of the coordinate system, whereas \( n \) and \( \lambda \) do not. Hence, since the physical oscillations of a spherically symmetrical system cannot depend on the coordinates in which they are described, the eigenfrequencies \( \omega \) cannot depend on \( m \).
The two sequences of modes are called \( p \) modes and \( g \) modes. The former have \( n > 1 \), and are standing acoustic waves. Pressure perturbations provide the dominant restoring force — hence the designation \( p \) modes. The \( g \) modes are essentially internal gravity modes, for which the dominant restoring force is buoyancy. For these \( n \leq 1 \). In addition, there is the so-called \( f \) mode (fundamental mode) which has \( n = 0 \). That mode has the character of a surface gravity mode, at least when \( \lambda \) is sufficiently large, and in that case there are no nodes in \( \zeta(r) \). Unfortunately this nomenclature has in the past led to some confusion, partly, perhaps, because when \( n \) and \( \lambda \) are of order unity the classification is not quite as straightforward as I have described: low-order low-degree modes can have the character of acoustic waves in one part of the star and of gravity waves in another. It is important that the \( f \) mode, which has little to do with acoustics, should not be confused with the fundamental (or gravest) \( p \) mode, which has \( n = 1 \).

**ACOUSTIC MODES OF HIGH DEGREE**

When \( \lambda \gg n \), the horizontal component of the wave number greatly exceeds the vertical component, and acoustic waves travel nearly horizontally. However, propagation is not precisely horizontal — locally horizontal propagation can never be sustained in an atmosphere with a vertical temperature gradient. Any wave front experiences a higher sound speed at greater depths, which refracts a descending wave back towards the photosphere. As it returns, the wave then experiences a background of decreasing scale height, and at about the level at which scale height and vertical wavelength are equal, the wave is reflected downwards: an oscillation cannot propagate through an atmosphere which varies more rapidly than the wave; instead the atmosphere is simply lifted coherently up and down. For an isothermal atmosphere the frequency \( \omega_c \) below which vertical propagation is impossible, which was discovered by Lamb (1909), is now called Lamb's critical (cutoff) frequency. The value of \( \omega_c \) is somewhat different for oblique waves. However, in and immediately below the solar atmosphere the scale height is much less than the horizontal wavelength of any wave that I shall discuss here. So when a wave approaches close to the photosphere, it is reflected where \( \omega_c = \omega \). At that point it is traveling nearly vertically, even though \( \lambda \gg n \).

High-degree \( p \) modes are thus confined within a shallow cavity beneath the photosphere. This acts as a wave guide, and determines the eigenfrequencies of the modes. In the sun the modes of greatest amplitude have frequencies near 3 mHz; they are the five-minute oscillations discovered by Leighton et al. (1962). Their dispersion relation, the relation between \( \omega \) and \( \lambda \), was first measured by Deubner (1975), who obtained a two-dimensional power spectrum in frequency and horizontal wave number \( k \) in the photosphere for waves traveling around the equator. [For such modes, \( k = R_{\odot}^{-1} \sqrt{m(m+1)} \) and \( m = \lambda \).] This has been compared with theory to calibrate the stratification of the solar convection zone.
The theoretical computations are numerical, and describe the physics as accurately as has been possible. However, the results can be easily understood from simple physical arguments. First notice that since the waves are confined quite close to the surface, the curvature of the sun is unimportant. Moreover, most of the confining cavity is adiabatically stratified; the temperature $T$ increases roughly linearly with depth $z$, and the sound speed $c$ is roughly proportional to $z^{1/2}$. Such functions define no natural length scale. Thus we deduce immediately that the only length scale that can influence the wave is the horizontal wavelength of the wave itself. Consequently, on dimensional grounds alone, we deduce that the depth $D$ of the acoustical cavity is proportional to $k^{-1}$, and does not depend on the actual value of the temperature gradient in the convection zone. The observations resolve the dispersion relation for $\ell$ between about 100 and 1000. Therefore to within factors of order unity, $D$ is no greater than about $0.01 \, R_\odot$, which justifies having neglected the sun's curvature. (Actually $D=2nk^{-1}$, and modes with $n \geq 10$ have been observed (Duval and Harvey, private communication). Therefore there is some information available about the outer 20 per cent of the solar radius.)

Turning now to Figure 1, one observes that in the outer layers $T$ (and hence $c$) is greater in Model B than in Model A. Therefore, since the penetration depths $D$ are the same for the two models, frequencies are greater for Model B. The observed dispersion relation can therefore be used to choose between the models.

Detailed studies of the sensitivity of the five-minute eigenfrequencies of high-degree modes have been made by Berthomieu et al. (1980) and Lubow et al. (1980). Because the modes do not penetrate beneath the convection zone, the theorists were content to study solar envelope models in isolation from the core. It was found that only with a relatively deep convection zone (about $2 \times 10^5$ km) could agreement with observation be obtained. This result is consistent with more recent, though apparently less thorough computations by Scuflaire et al. (1981). Thus, as is evident from Figure 1, the calibration implies a relatively high helium abundance, a high central temperature and consequently a high neutrino flux. Note, however, that the deduction about conditions in the core demanded an application of the theory of stellar evolution, and therefore depends on the theory's validity.

THE F MODES OF HIGH DEGREE

In addition to the acoustic modes, the power spectra of the observations of solar oscillations of high degree contain $f$ modes. It is easy to show that these have the dispersion relation $\omega^2 = gk$; the modes are solenoidal and irrotational, with $\xi, \eta = e^{-kz}$. This result is valid in the limit $kr_0^{-1} = \ell \to \infty$, and is essentially independent of the stratification of the envelope. Thus, the high-degree $f$ modes have little value for diagnosing the mean stratification of the envelope. But they are of use. Since we know that the $f$-mode frequencies of all solar models must converge to $(gk)^{1/2}$ as $k$
Increases [with a relative deviation from that limit of about $\chi^{-1}$], the $f$-mode eigenvalues may be used to assess the errors in published numerical studies.

**ACOUSTIC MODES OF LOW DEGREE**

When $n >> \ell$, propagation is nearly vertical, and the modes penetrate to the core of the sun. These oscillations therefore provide a direct probe of the deep interior. They have been observed in the sun principally in the frequency range 2-4 mHz (Scherrer 1982).

When $n$ is large the length scale of the density perturbation is small. Therefore the magnitude of the perturbation $\Phi'$ to the gravitational potential is small too. Consequently, in asymptotic studies at large $n$ (e.g. Vandakurov 1957, Tassoul 1980), $\Phi'$ has been ignored. The resulting dispersion relation is of the form

$$v = \nu_0 \left(n + \frac{1}{2} \ell + n_e\right) - A[\ell(\ell+1) + \delta] \nu_0^{-1} + \ldots,$$

where $v = \omega/2\pi$ is the cyclic eigenfrequency,

$$\nu_0 = \left(2 \int_0^\infty c^{-1} \, dr \right)^{-1},$$

$n_e$ is an effective polytropic index of the outer layers of the envelope (in the vicinity of the reflecting layer, where $\omega_e = \omega$), and $A$ and $\delta$ are more complicated integrals of the equilibrium model.

The functional form of the dispersion relation (4) is a geometrical property. It is the same as the dispersion relation for sound waves in an isothermal homogeneous gas contained in a rigid spherical cavity. In that case the coefficients $A$ and $\delta$ are purely geometrical quantities, and are given by the expansion of the high-order zeros of a spherical Bessel function. In an inhomogeneous sphere such as the sun, these quantities depend also on the stratification.

The observations of low-degree modes use techniques with little or no spatial resolution (see discussions by Scherrer and Christensen-Dalsgaard in these proceedings). The degrees of the high-order acoustic modes, for example, are inferred by comparing with observation the pattern in which the eigenfrequencies (4) and the anticipated amplitudes are distributed (see Christensen-Dalsgaard 1982). Since the signal amplitudes depend strongly on the structure of the spherical harmonics in the eigenfunctions, they, like the distribution of frequencies, reflect mainly the geometry of a sphere. Thus the identification of $\ell$ is not sensitive to the detailed structure of the solar model, and is therefore robust. The identification of $n$, on the other hand, is model-dependent. Since one cannot yet count down to the fundamental acoustic mode in the data, one must rely on a comparison of the magnitudes of the frequencies observed with the values of theoretical eigenfrequencies. In
so doing, one can select a preferred solar model. As is evident from Eq. (4), the eigenfrequencies are determined mainly by $v_0$, which is related to the sound travel time from center to surface. That is controlled mainly by temperature, and since Model A is hotter than Model B throughout most of the interior (see Fig. 1), it has higher frequencies (Iben and Mahaffy 1976, Christensen-Dalsgaard et al. 1979). As $Y$ varies, the frequencies of the five-minute modes of low degree change in the opposite sense to those of the modes of high degree.

A least-squares fit of all the available low-degree five-minute data with eigenfrequencies $v_\lambda$ linearly interpolated between or extrapolated from Models A and B according to

$$v_\lambda = \lambda v_A + (1-\lambda)v_B$$

(where $v_A$ and $v_B$ are corresponding eigenfrequencies of Models A and B) yields $\lambda = 1.36 \pm 0.12$ (Christensen-Dalsgaard and Gough 1981). If a similar interpolation in $Y$ were valid, this would imply $Y = 0.27 \pm 0.01$. There is a second possibility that $\lambda = -0.32 \pm 0.08$, but the fit with observation is not as good. The difference between the quality of the two fits in the frequency range in which the observations overlap is not much greater than the discrepancies between the different observations, so the preference for the high-$\lambda$ fit on the basis of low-degree modes alone is not overwhelming. However, only the higher value of $\lambda$ is consistent with the calibration by the five-minute modes of high degree. Notice, however, that whereas the high-degree modes measure only the structure in the upper part of the convection zone, the low-degree modes measure an integral property of the entire sun. Therefore the two calibrations are rather different, and their consistency is evidence that standard solar models are broadly correct.

There is evidence, however, that the models are not quite correct, for there are systematic deviations between the best-fitting frequencies $v_\lambda$ and reality. Most apparent is that the mean spacing between theoretical frequencies of adjacent order and like degree is too great. This deficiency was found in several models of Christensen-Dalsgaard and Gough (1980), some of whose calculations were confirmed by Shibahashi and Osaki (1981). (There are also some computations by Scuflaire et al. (1981), but since these did not use complete solar models the conclusions concerning low-degree modes must be treated with some caution.)

In trying to understand why the theoretical eigenfrequencies are not correct, it is instructive to look in more detail at the structure of Eq. (4). The order $n$ is large (typically between about 15 and 30 for the observations currently at hand) and $\ell$ is relatively small ($\ell \leq 5$). The dimensionless constants $n_\text{e}$, $A$ and $\delta$ have magnitudes of order unity (in the sense that $\ell(\ell+1)$ is of order unity!), and the second term in Eq. (4) is a small correction to the first. Therefore the eigenfrequencies lie in approximately uniformly spaced groups, of alternately odd and even degree, and which can be labeled by the parameter $x \equiv n + \ell/2$. The mean separation is roughly $v_0/2$, and the curvature
of \( v(x) \) depends on \( A \) and \( \delta \); the absolute values of the frequencies depend also on the constant \( n_e \).

If, to the degree of accuracy in which we are interested, Eq. (4) were a good approximation, one could gain some insight by separating \( v(x,\lambda) \) into its \( \lambda \)-dependent and \( \lambda \)-independent parts. Since sound takes 2 hours to travel from one side of the sun to the other, purely acoustic oscillations cannot interact throughout the envelope during a single period. Therefore in the envelope the dynamics of the low-degree oscillations should not sense the horizontal wave pattern; the eigenfunctions are essentially independent of \( \lambda \) (aside from the dependence through \( x \)). Only in the core, at radii \( r \) that satisfy \( r/\lambda c \ll v^{-1} \), does the motion sense the horizontal variation of the mode. Thus one might infer that the \( \lambda \)-independent part of the discrepancy between observation and theory would control the revision that is required in the envelope of the solar model, and the \( \lambda \)-dependent part would control the core. In particular, Dziembowski and Gough (unpublished) have shown how \( \lambda(\lambda+1)A \) in Eq. (4) would be related to the curvature of the sound speed in the core. The hope was that this might lead directly to a determination of whether the material processed by the nuclear reactions has been mixed with its environment, or whether thermal balance has possibly been upset.

Unfortunately, it transpired that so simple a description is inadequate: the error introduced by neglecting the perturbation \( \Phi' \) to the gravitational potential is comparable with the \( \lambda(\lambda+1) \)-dependent term in Eq. (4). A generalization of Tassoul's (1980) analysis to include \( \Phi' \) is therefore necessary, and indeed is now under way.

**ON SOLAR INVERSION WITH LOW-DEGREE MODES**

Since the low-degree \( p \) modes penetrate to the core, it is interesting to inquire whether they are sufficient to measure the entire stratification of the sun. I report here the results of some unpublished work by A. J. Cooper and myself that throw light on this issue. The aim was to discover what can be inferred about the structure of a simple solar model from a knowledge of some of its oscillation eigenfrequencies.

Two simple solar models were constructed: Model S, which we pretend is the real Sun, and Model T, which is a trial. Both models are composite polytropes, with index 3 out to a radius \( r_c \) and index 3/2 beyond. Thus they each represent a star with a radiative interior and an outer convection zone. Both models were chosen to have the mass and radius of the sun, and in computing the oscillations the adiabatic exponent was taken to be 5/3. The only difference between the models was the value of \( r_c \): in Model S it was 0.75 and in Model T it was 0.80 [the calculations were performed before the calibrations of Berthomieu et al. 1980, Lubow et al. 1980 and Christensen-Dalsgaard and Gough 1981, when it was believed that \( r_c = 0.75 \) was a good estimate of the radius of the base of the convection zone.] Linearized adiabatic oscillation eigenmodes of both models were computed, and the frequencies of Model S were
regarded as artificial data. Two inversion procedures were then applied to
the differences in the eigenfrequencies of the two models in an attempt to
infer the structure of Model S.

Both procedures rely on the assumption that Model T is a good guess at
Model S. One starts from a variational principle (Chandrasekhar 1964), in
which $\omega$ is expressed as a functional of the structure of the equilibrium model
and the displacement eigenfunction $\varphi$, that functional being stationary with
respect to variations in $\varphi$ when the structure is fixed. One can then compute
the small deviations $\delta \omega$ from $\omega$ produced by small variations in the structure,
retaining only linear terms. Since the functional representation of $\omega$ is
stationary with respect to $\varphi$, the perturbations to $\varphi$ resulting from changing
the equilibrium model make only a quadratic contribution to $\delta \omega$, and can
therefore be ignored. After some straightforward manipulations, and use of
the equation of hydrostatic support and mass conservation to relate the
perturbations in the structure variables, the frequency change can be written
in the form

$$\frac{\delta \omega}{\omega} = \int_0^R F(\varphi; \rho) \frac{\delta \rho}{\rho} \, dr,$$

where $\rho$ is the density of the equilibrium model, and $\delta \rho$ is a small perturba-
tion from it. Thus if $\rho_S$ and $\rho_T$ are the densities of Models S and T, $\omega_S$
and $\omega_T$ are corresponding eigenfrequencies, and Model T is close to Model S,

$$\frac{\omega_T - \omega_S}{\omega_T + \omega_S} = \frac{1}{2} \int_0^R F(\varphi; \rho_T) \frac{\delta \rho}{\rho} \, dr,$$

for each eigenvalue, where now

$$\frac{\delta \rho}{\rho} = 2 \frac{\rho_T - \rho_S}{\rho_T + \rho_S},$$

and $\varphi$ is the eigenfunction of Model T.

Since $\omega_S$ is known (presumed measured), and so, of course, are $\rho_T$ and $\omega_T$,
it is possible to compute the kernel $F$ and regard Eq. (8) as an integral con-
straint on $\delta \rho/\rho$. From a variety of oscillation modes, one deduces a variety
of such constraints, with different kernels, from which it is hoped that the
function $\delta \rho/\rho$ can be estimated. [A selection of kernels is displayed by Gough
(1978).] From that one estimates $\rho_S$ from Eq. (9).

It is never possible to deduce $\delta \rho$ precisely. Part of the reason for that
is that to define a function requires an infinite amount of spatial informa-
tion, and we can never apply more than a finite number of integral con-
straints. Therefore one must adopt some criterion for choosing a solution.
Two separate criteria have been chosen.

Since we assumed at the outset that we were good at guessing the model:
namely that Model T was in some sense close to Model S, we can take this
prejudice to its logical conclusion and seek the model that is in some sense closest to Model T but which has the 'observed' eigenfrequencies \( \omega_S \). This we do by minimizing

\[
\int_0^R W(r)(\delta \rho/\rho)^2 \, dr
\]

for some weight function \( W(r) \) amongst all functions \( \delta \rho \) satisfying

\[
\int_0^R r^2 \delta \rho \, dr = 0
\]

(to preserve total mass) and the constraints (8). Having computed an estimate of \( \rho_S \) from Eq. (9), one can iterate the entire procedure using that estimate as the new trial. We have found that this iteration can converge, depending on the modes selected for the constraints (8), but that the resulting estimate of \( \rho_S \) can develop small-scale structure which is evidently artificial; a smoothness criterion should perhaps be incorporated into the minimization procedure. Therefore I shall report here only on the results of the first iteration.

The second method we adopted is the optimal averaging procedure of Backus and Gilbert (1970). Here linear combinations of the constraints (8) are chosen such that the corresponding combination of kernels \( F \) resembles a delta function at \( r = r_0 \). The delta-function-like combination is called the optimal averaging kernel. The integral is then approximated by \( \rho(r_0) \), which is thus determined as a linear combination of the data. By repeating the procedure for different values of \( r_0 \) one can construct \( \rho(r) \).

One advantage of the optimal averaging procedure is that one can inspect the optimal averaging kernels to see how broad they are, and thus estimate the scale of resolution of \( \rho(r) \). Another, related advantage is that there are sometimes ranges of \( r_0 \) at which the procedure fails; this is informative, because it implies that the data provide no useful information about \( \rho(r) \) in those ranges. [The first method of inversion has the disadvantage of always producing a result.] However, the procedure does not produce a model that satisfies the constraints (8) precisely. [One could, of course, use this result as a trial for the first method, and hence find the closest model to it that does satisfy the constraints.]

The optimal averaging procedure can be iterated upon too. But, in contrast to the previous method, the iterated solution, when it is converged, does not depend upon the initial trial model [though, of course, if several solutions exist, the one to which the iterations converge does depend on the trial.] The solution is biased only by being smooth, since averaging tends to smooth a function. The degree of smoothness is such that the solution contains hardly any small-scale structure that is impossible to resolve with the data.
The purpose of our investigation was to test methods of inversion and to assess the information content of various combinations of modes. Since our data were constructed from a known model, we knew the correct answer in advance, and could therefore measure the accuracy of our methods. Our tests, which are based on linear theory, were perhaps somewhat severe, because the true value of $\delta p/p$ relating Models S and T is not everywhere small, and in places is as great as 40 per cent.

First we choose two sequences of p modes, with $\ell = 1$ and $\ell = 2$ and with $1 \leq n \leq 10$. Inversion by both methods was rather poor. In places the values of $\delta p/p$ deduced were in error by 50 per cent of the true rms value. However, it was interesting to note that the regions where the optimal averaging procedure failed to produce a solution were more or less the same as those regions where the first method gave the least accurate results. Interestingly, those regions were $0 \leq r/R_\odot \leq 0.25$, $0.65 \leq r/R_\odot \leq 0.85$ and $r/R_\odot \geq 0.97$. Thus the low-degree p modes of moderate order give information about the outer layers of the sun (excluding the outer few per cent — but that is well measured by the high-degree modes) and a region at the base of the envelope. One would anticipate that the core could be penetrated by increasing $n$, but we have not yet tried that exercise. Replacing the $\ell = 2$ sequence by a similar $\ell = 0$ sequence gave comparable results.

The second experiment was to add to the data the $f(\ell=2)$ mode and some low-order g modes. This was done because it was estimated that such modes would be detected by a full-disk measurement from space. The result was that now both methods give results accurate to just a few per cent over almost the entire range of $r$.

One should not conclude immediately that all that is now needed for measuring the mass distribution in the sun is the frequencies of the low-degree p modes and a few g modes. The experiments reported here used very simple models, and though they did not rely on the fact that Model S is a composite polytrope, the inversion procedures may turn out to be less reliable with more realistic physics and real data. One cannot conclude either that the g modes are essential. It may be that sufficient information is contained in the p modes of intermediate $\ell$. This has yet to be tested. Nevertheless the experiment does present us with the exciting prospect of being able to measure the solar interior in the near future.

ON THE SEAT OF THE SOLAR CYCLE

I shall not discuss the dynamics of the solar cycle. I intend merely to point out that the frequencies of the oscillations I have been discussing must respond to changes in the structure of the sun, and that if these are detectable they might be useful for diagnosing the nature of the cycle.

There has been some discussion about the response of the sun's structure to hydrostatic perturbations associated with the cycle (e.g. Sofia 1981). In
particular, there is limited evidence that the sun's luminosity should vary by perhaps a few tenths per cent. The radius should vary too, though to a rather lesser extent. The ratio of the relative variations in radius and luminosity is a diagnostic of the dynamics, and there is a hint that its value should be some indication of the mean depth of the major processes that perturb the structure (Gough 1981a).

To date, little work has been carried out on the potential variations of the oscillation eigenfrequencies. However, there is some evidence that there should be measurable changes to the p modes of low degree. This comes from calculating the change in $v_o$, defined by Eq. (5). Though $v_o$ alone does not determine the frequencies accurately, even when $n$ is large, Christensen-Dalsgaard and I (unpublished) have found that when, for example, the chemical composition of the equilibrium model is changed, the variations in $v_o$ reflect quite well the variations in the computed eigenfrequencies. So there is reason to suspect that the same is true of simulated solar-cycle variations. Thus it can be inferred, for example, that if luminosity variations were produced solely by the blocking by sunspots of radiation in the immediately subphotospheric layers, a process that has been frequently modeled in recent years by varying the mixing length, then it is unlikely that the frequencies of the low-degree five-minute modes would change by much more than a part in $10^5$. But if, on the other hand, 0.1 per cent variations were produced deeper down, say by magnetic suppression at the base of the convection zone, as has been suggested by Spiegel and Weiss, a change as large as 1.5 μHz or so in the frequencies could result. The precise magnitude of the change depends on how the thermal stratification relaxes in the inhibited layers. The phase is such that the eigenfrequencies should be greatest at sunspot maximum. Such changes should be detectable, but at present they are smaller than the discrepancies between the results of different observers (cf. Christensen-Dalsgaard 1982, Scherrer et al. 1982).

Magnetic suppression of convection at the base of the zone could also change the low-degree g mode frequencies. Once again, the changes depend on the thermal readjustments, and also on the direct effect of the Lorentz forces on the oscillations. Variations of order 1 μHz are not out of the question, but in this case, if thermal perturbations dominate the change, the extension of the radiative region is likely to yield greatest frequencies at sunspot minimum.

It is quite unlikely that, to current observational accuracy, the solar cycle will produce perceptible variations in the mean positions of the ridges in the k-ω spectra of the p modes of high degree. There might, however, be measurable changes in the variations of the ridges caused by changes associated with the solar cycle in the structure of the large-scale convective flow.
GIANT CELLS

Inhomogeneities in the solar envelope associated with convection modify the simple oscillation eigenvalue calculations that I have been discussing. When the convection length scale is large compared with the horizontal wavelength of the mode, as is the case with high-degree p modes on giant cells, the acoustic wave experiences a slowly varying background, and responds locally. Perturbations in c and γ (and, to some degree, vertical velocities) change the absolute dispersion relation, and horizontal velocities in the direction of the horizontal wave number advect the wave pattern. The two processes have different symmetries with regard to corresponding waves with positive and negative phase speed, and can therefore be separated.

A preliminary analysis of the distortions of the wave pattern produced by convection in a plane polytropic layer has been presented by Gough and Toomre (1982). Observational evidence for such distortions has been reported by F. Hill et al. (1982a and in these proceedings). An inversion analysis of the data is currently in progress.

ROTATIONAL SPLITTING

By the combined influence of advection and Coriolis forces, slow rotation about a fixed axis causes standing wave patterns to precess about that axis. This is manifest as a splitting of the degeneracy with respect to m of the component propagating waves. Here m is defined with respect to spherical polar coordinates about the rotation axis. The cyclic frequencies thus become $v_{nlm} = v_n + v'_{nlm}$, where $v_n(\pi, \theta)$ are the frequencies in the absence of rotation and $v'_{nlm}$ can be written as a linear functional of the angular velocity $\Omega$:

$$v'_{nlm} = m \int K_{nlm}(r, \theta)\Omega(r, \theta)r^2 \sin \theta \, dr \, d\theta$$

(12)

Thus by measuring the rotational perturbations $v'_{nlm}$ for a variety of modes, $\Omega$ can be estimated by inversion techniques.

The kernels $K_{nlm}$ have the property that if $\Omega$ is a function of $r$ alone, the integral in Eq. (12) is independent of $m$. In that case $v'_{nlm}$ is proportional to $m$, there is no azimuthal dispersion, and wave patterns of given order and degree precess without change of shape.

Notice that of the two inversion techniques discussed earlier, the first cannot be used unless one has a preconceived idea of what $\Omega$ should be. However, it is straightforward to apply the optimal averaging procedure. Recall that the resulting estimate of $\Omega$ is actually an average of the true $\Omega$, assuming the data to be correct, and is therefore likely to be smoother than the true $\Omega$.

There are reports of direct evidence for rotational splitting in two sets of data: the full-disk Doppler measurements of Claverie et al. (1981) and
limb-darkening measurements by Bos and Hill (reported by Hill 1982). It has also been suggested that rotational splitting is responsible for the time variations in the Princeton oblateness data (Gough 1981b). Since there is some uncertainty in the interpretation of these data, I shall first describe briefly what the evidence is.

Peaks near 3 mHz in the power spectrum of recent full-disk Doppler data obtained by Claverie et al. (1981) were found to be multiple. It was difficult to resolve the structure in individual peaks, so to reduce noise a superposed frequency analysis was performed. The cleanest results were obtained by aligning the $\ell = 0$ modes, which should not be multiple. It was found that the $\ell = 1$ peak was split into three components of approximately equal power, and the $\ell = 2$ peak appeared to be split into five. These were interpreted by Claverie et al. as the $2\ell+1$ rotationally split modes, whose mean spacing was measured to be 0.75 $\mu$Hz.

The immediately obvious objection to this interpretation is that one should not see all $2\ell+1$ peaks. As Claverie et al. acknowledge, the symmetry of some of the modes precludes their detection by full-disk observations. Indeed, any circularly symmetrical observation of the solar disk cannot detect modes for which $\ell + m$ is odd, provided the rotation axis is perpendicular to the line of sight (Christensen-Dalsgaard and Gough 1982, Christensen-Dalsgaard 1982). Actually, at the time of the observations there was a 6$^\circ$ tilt of the axis, but that cannot explain the observations unless modes with odd $\ell + m$ systematically have greater power than the others. Therefore the claim that low-degree five-minute modes are rotationally split by 0.75 $\mu$Hz should be regarded with some caution.

Limb-darkening data were obtained by Bos and Hill (1982) on 18 days during a 41-day period. The spectrum of the entire discontinuous data set was studied in the frequency range 250–800 $\mu$Hz. This range was chosen because it includes the maximum of the buoyancy frequency in standard solar models. Below that frequency the $g$ modes are densely spaced, but above it there can be only $p$ modes; moreover below 800 $\mu$Hz the latter are sparsely distributed. Therefore there should be a good chance of isolating the modes.

The spectrum is composed of approximately uniformly spaced peaks with a mean separation of roughly 0.6 $\mu$Hz. The peak widths are consistent with the 0.29 $\mu$Hz natural width corresponding to the 41-day observing period. Bos and Hill (1982) and H. A. Hill et al. (1982a) have argued that essentially all the power is solar. Interestingly, there seems to be no evidence for a change in either the power or the spacing of the peaks where the $g$-mode spectrum is expected to terminate. But this may be so simply because interference between unresolved more-densely spaced modes is likely to produce peaks with a mean separation roughly twice their natural width.

In an attempt to isolate the modes in the spectrum, groups of uniformly spaced peaks were sought (Hill 1982, H. A. Hill et al. 1982b, Gough 1982a). This is what one would expect of rotational splitting if $\Omega$ were a function of
r alone. Seven groups were found and, taking into account the modes that should not be detectable because of their spatial symmetry, a lower bound to $\lambda$ was deduced. The inferred rotational splitting varied between 0.74 and 2.90 $\mu$Hz. The rotationally split modes, if that is what they are, were then identified by comparing their frequencies with theory.

I should point out an interesting feature of this procedure. The identification is based on the expectation that uniformly-spaced peaks are unlikely to occur accidentally. Indeed, to be reasonably confident that that is so, one must accept only groups with many components, which precludes isolating modes of very low degree by this method alone. However, the result must necessarily be that the peaks found are uniformly spaced, and if these are assumed to be rotationally split modes, one deduces that $\varpi$ is essentially a function of $r$ alone. To be consistent with the data the latitudinal dependence of $\varpi$ in the interior of the sun must be much less than in the photosphere. Indeed, if the photospheric differential rotation persists to a depth $d$, and beneath that $\varpi$ is independent of latitude, $d$ would be only about $0.05 \, R_\odot$. If that is really so, a severe constraint on dynamical theories of the convection zone will be imposed.

The implications of the claims for rotational splitting have been discussed by H. A. Hill et al. (1982b) and Gough (1982a). Using the optimal averaging procedure of Backus and Gilbert (1970), Gough obtained an estimate of the angular velocity $\varpi$ that varied smoothly from about $1.9 \times 10^{-5}$ s$^{-1}$ at the center to the observed equatorial value of $2.9 \times 10^{-6}$ s$^{-1}$ at the surface. If this represents the true solar rotation, significant material mixing by the Eddington-Sweet circulation might have taken place. It was found also that the widths of the optimal averaging kernels were large, and the details of $\varpi$ correspondingly poorly determined, in the mid-range: $0.2 \lesssim r/R_\odot \lesssim 0.8$. This is the region that contributes the most to the quadrupole moment $J_2$ of the solar gravitational field, a quantity of crucial importance for testing theories of gravitation using the orbits of planets.

There has been some dispute about the possible values of $J_2$ that are consistent with the data. If one applies no constraints other than those imposed by the apparent splitting data, one can deduce only that $J_2 \gtrsim 1.2 \times 10^{-6}$ (Gough 1982a). This is well within the limits set by analyses of the orbit of Mercury, if General Relativity is correct. On the other hand, H. A. Hill et al. (1982b) claim that $J_2 = (5.7 \pm 1.5) \times 10^{-6}$, which is two standard deviations away; this conclusion relies on their having restricted $\varpi$ to monotonic decreasing curves "with a minimum structure" (which they do not define). The optimal averaging procedure yields $J_2 = 3.6 \times 10^{-6}$, which deviates from the orbit data by just one standard deviation if General Relativity is correct.

A prominent property of all functions $\varpi(r)$ that satisfy these rotational splitting constraints is that somewhere they must decrease with $r$ faster than $r^{-2}$. This implies that they are unstable to diffusive instabilities (Goldreich and Schubert 1967, Fricke 1968). That is not to say that the data have necessarily been interpreted incorrectly, because there may be angular-
momentum transporting processes, associated perhaps with waves or magnetic fields, to maintain the high rotation gradient against the natural tendency of the instabilities to destroy it. Nevertheless, if that possibility is accepted, in our present state of knowledge one cannot maintain the commonly held view that \( \Omega(r) \) is necessarily monotonic. With regard to this issue it is interesting to note that after taking account of the solar rotation the phase coherence of diameter measurements reported previously by H. A. Hill and his colleagues is greater than one would expect from modes with infinite lifetimes and randomly distributed amplitudes and phases (Christensen-Dalsgard and Gough, unpublished; cf. discussions by Hill and Caudell 1979, Caudell et al. 1980 and Gough 1980). This is evidence that there is a systematic preference for either eastward or westward propagating waves in any group of modes of like order and degree [though the preferred direction could be different for different groups], a property which is essential for angular-momentum transport by waves.

THE DIRECT INFLUENCE OF THE MAGNETIC FIELD ON SOLAR OSCILLATIONS

From an analysis of the temporal variations in the Princeton oblateness data Dicke (1982) has inferred the existence of an intense magnetic field in the sun, extending out to a radius of about 0.95 \( R_\odot \). [Interestingly, this is the same as the value one might deduce from the data of Bos and Hill (1982) for the radius below which differential rotation might cease.] Aside from problems concerning the convection zone that this raises, Dicke's field configuration is probably not inconsistent with the theoretical studies by Mestel and his collaborators (e.g. Mestel and Takhar 1972). Isaak (1982) claims that the existence of all 2\( n+1 \) peaks in the power spectrum of the Birmingham full-disk Doppler data supports Dicke's contention, but it is not clear why that should be so (Gough 1982b).

In the past, the possibility of a measurable modification of the eigenfrequencies by an intense magnetic field has received little attention. In the future, it is likely to be taken more seriously.

Support by the Science and Engineering Research Council, the National Aeronautics and Space Administration through grant NSG-7511 and the Air Force Geophysics Laboratory through contract F19628-77-C-0104 to the University of Colorado is gratefully acknowledged. The Joint Institute for Laboratory Astrophysics is operated jointly by the University of Colorado and the National Bureau of Standards.

REFERENCES

Backus, G. and Gilbert, F. 1970, Phil. Trans., A266, 123.
Christensen-Dalsgaard, J. 1982, these proceedings.
Hill, F., Toomre, J. and November, L. J. 1982b, these proceedings.
Scherrer, P. H. 1982, these proceedings.