44

DOUGLAS O. GOUGH* AND PHILIP H. SCHERRER**

The solar interior

The quest to learn the internal structure and dynamics of the Sun is motivated by several issues. For most people the most obvious is the desire to understand the source of energy for the Earth, which is essential for the maintenance of life. Astronomers are interested in the Sun because it is an example of a typical main-sequence star that can be studied in enormously greater detail than any other. And in addition, the Sun can be used to investigate fundamental physics: it is an important source of gravity, providing a testbed for the general theory of relativity, and it contains material at high pressure and temperature permitting us to study particle, nuclear and atomic physics, plasma physics, and fluid dynamics under conditions that cannot be achieved on Earth. Perhaps less fundamental but certainly more important to society, the solar interior is the source of both secular and cyclic variability in the electromagnetic and particle fluxes, and of all their effects on the Earth and human technological systems.

This chapter describes first our understanding of the solar interior prior to 1975. It then provides a narrative discussion of developments in each of several major topics since that time. These topics include neutrinos, shape, irradiance, calibration of stellar models, rotation, the dynamics of the convection zone and of large-scale features within it, magnetic field generation and the solar cycle, and the nature of active regions and active longitudes. Beginning in about 1975 a new tool for probing the interior of the Sun, and some day of other stars, was discovered. This tool is helioseismology. The impact of helioseismological inferences on our understanding of the solar interior is so dominant that a brief review of the techniques by which it is used will be provided prior to an examination of the development and present state of each of the above topics. For the most part the further development of the main scientific topics has followed the development of helioseismology, and therefore the discussion of those topics will be intertwined with the discussion of helioseismology.

STATUS PRIOR TO 1975

Before 1975 the only observable quantities that yield information about the solar interior were the global parameters – mass, radius, luminosity, and effective temperature – which were used to calibrate (spherically symmetrical) theoretical models. A measure of the neutrino flux, which reveals energy generation processes directly, was also available. The age of the Sun is inferred from the age of the Earth, and other interplanetary material. The observed surface differential rotation and the existence and organization of magnetic fields in the photosphere also provided information about processes assumed to be operating in the convection zone. There was also a measure of the oblateness of the photosphere, from which one should be able to infer a measure of the mean rotation of the interior and the oblateness of the exterior gravitational equipotentials. But the interpretation of some of those data was in question. Primarily, our understanding of the solar interior was based on the application of stellar models.

Prior to the observation that there were too few neutrinos coming from the Sun, theoretical models could be adjusted to match the global observables. Explanation of the low neutrino flux was a big problem. Although in principle the solar data were adequate to determine the uncertain parameters that specify the simplest theoretical models – the initial helium abundance Y, the total heavy-element abundance Z (the relative proportions of most of the heavy elements can be determined by spectroscopic analysis) and a scaling parameter α contained in the mixing-length theory used to model the convective heat flux – the outcome was

* University of Cambridge, United Kingdom
** Stanford University, CA, USA

Douglas O. Gough and Philip H. Scherrer, The Century of Space Science, 1035–1063

generally not believed, for it yielded an implausibly low value of the helium abundance, namely $Y \simeq 0.16$. Even though it was in the Sun that helium was discovered, a direct abundance determination is not possible. The atmosphere is too cool for a spectroscopic analysis, and the theory of radiative transfer in the corona is too uncertain to yield a reliable value. Abundance measurements were available from *in situ* measurements of the solar wind, but these exhibited temporal and spatial variability, implying that chemical differentiation had taken place and that therefore none of the measured values necessarily represents the value in the Sun. However, there were astronomical estimates: values of Y determined in the atmospheres of hot stars believed to be of comparable age to the Sun and measurements in ionized interstellar gas clouds suggested a value of about 0.25. Moreover, calibrated theories of Big Bang nucleosynthesis yielded values of Y that exceeded 0.20 – comfortably less than the values observed in the stars which have condensed from gas clouds that have been enriched by nuclear processed material from supernova explosions, yet substantially greater than the value from the solar calibration. Therefore the value of 0.16 was not generally accepted; instead it was assumed that $Y \simeq 0.25$, and that it was the neutrino flux that could not be explained. Consequently the issue was dubbed "the solar neutrino problem." Nevertheless, there remained some doubt about Y.

MOST IMPORTANT ISSUES IN 1975

(i) Solar neutrino flux: Is it a problem with the structure of the Sun, or with nuclear physics or even particle physics? Spontaneous oscillations between neutrino flavors had been suggested by Pontecorvo (1968), raising the possibility that a substantial fraction of the electron neutrinos that are produced in the Sun have been transformed into μ or τ neutrinos by the time they reach the Earth, and thereby had evaded detection. However, for such a process to occur, neutrinos would need to have mass, and at the time it was almost universally believed that neutrinos were massless. There had been a great industry in adjusting parameters of standard solar models, and adding extra ingredients almost always in a spherically symmetrical fashion. That some neutrinos were detected at all was regarded as a triumph for nuclear physics. A crucial assumption in the construction of the standard models is that they are in thermal balance (the thermal relaxation time is very much less than the characteristic time of structural variations arising from chemical composition changes produced by nuclear transmutations). This implies that the integrated rate of generation of thermal energy by nuclear reactions is equal to the "observed" luminosity of the Sun (at least, the luminosity inferred from irradiance measurements, assuming spherical symmetry). A second assumption is that the core is motionless, and that therefore all but the slowest nuclear reactions have reached local equilibrium. Together, these assumptions impose tight theoretical constraints on the balance of reactions and the consequent neutrino production rates.

(ii) Oblateness of gravitational equipotentials and the history of solar spin-down: How much greater is the angular velocity Ω in the core than it is at the surface? There had been many studies of potential instabilities arising from shear presumed to be imposed by spin-down and to which the Sun was regarded as being neutrally stable. They all implied that Ω increases inwards, which would cause the oblateness of the Sun to be greater than one might suppose from the surface rotation alone. Dicke and Goldenberg (1967, 1974) had claimed such a greater oblateness which might have been compatible with some of these theoretical studies, but there were problems with interpreting the optical measurements of the shape of the Sun, owing to there being a greater emission of light from the solar atmosphere in equatorial regions making the solar disk appear to be more oblate than the matter distribution. Pole–equator variation in convective flow could also influence the measurement. The contribution to the oblateness from the centrifugal force due to the rotation of the photosphere needs to be subtracted from the raw measurement: the residual would be only about 4% of the total if Ω were uniform, rendering the corrected measurement uncertain. A measurement of the oblateness of the surface would have been fine if Ω were much larger in the interior, as Dicke had predicted to be required for the Brans–Dicke theory of gravity.

The interpretation of the measurements of Dicke and Goldenberg were doubted by the community, partly because they contradicted general relativity by implying too great a precession of the perihelion of the orbit of Mercury. A subsequent contradictory measurement by Hill and Stebbins (1975) was accepted immediately by most astronomers, however, perhaps not as critically as one should expect. Nevertheless, it still appears that any measurement of the shape of the photosphere is an extremely unreliable guide to the oblateness of the gravitational equipotentials, and that more direct methods are preferable. As we describe below, the inference from helioseismology is now by far the best, and is likely to remain so for a long time.

(iii) The source of magnetic fields and the activity cycle was also a mystery. While a number of competing models were put forward, the general belief was that a dynamo process operating in the convection zone and driven by rotational shear is the source of active-region fields and of the 22-year magnetic cycle. A phenomenological discussion invoking differential rotation, buoyancy, and supergranulation to distribute the fields into the observed patterns (Babcock 1961, Leighton 1969) has held considerable credence, and it is not unlikely that these processes form the

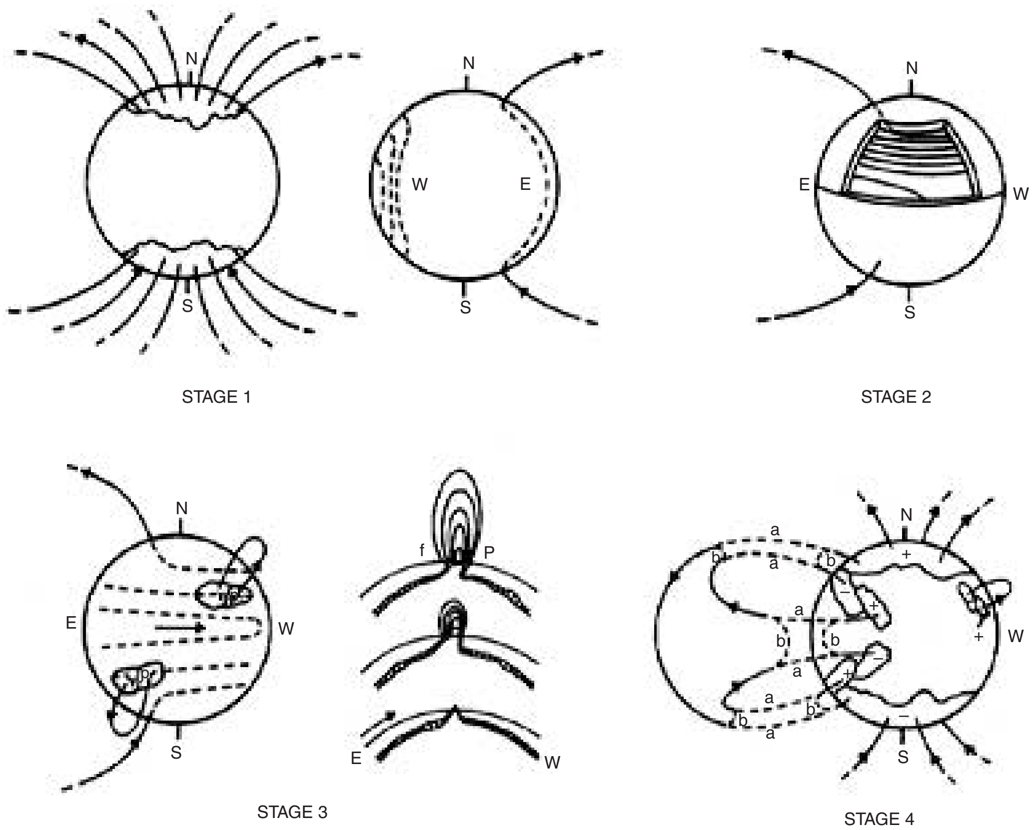

Figure 1 Schematic from Babcock (1961) via Rabin *et al.* (1991) showing the proposed evolution of the 22-year solar magnetic cycle, which still forms the basis of many modern views. Stage 1 is a state of basically poloidal field. In stage 2 the field lines are stretched into toroidal spirals by the differential rotation (principally in the tachocline, perhaps), and once the field has reached a strength sufficient to become buoyant, parcels of fluid rise through the convection zone drawing with them loops of field which erupt through the photosphere (stage 3) to form bipolar magnetic regions. In stage 4 advection by a combination of differential rotation and the (anisotropic) convection causes leading active regions to migrate equatorward and trailing regions to migrate poleward, creating new poloidal field of opposite polarity; the toroidal field dissipates to produce stage 5, which is like stage 1 but with the sense of the field reversed.

basis of much of what is happening. Figure 1 shows the essential characteristics of these models. But there was no precise model dynamo that could reproduce the observed magnetic cycle as shown in Figure 2. The observed fact of "active longitudes" or "zones" could not be accounted for in any model. Much progress had been made to understand the source of coronal and interplanetary fields in terms of the photospheric fields, but there was no understanding of the origin of the photospheric fields, nor of their organization.

(iv) The stability of the "solar constant," or total irradiance was unclear. There was historical evidence of secular changes, and there was uncertainty about the variability with activity. The correlation of solar-activity indices with northern European climate records suggested strongly that there is at least a long-term coupling of solar activity with climate (Eddy 1976). Attempts to measure total irradiance and its variations from the ground, under way since the beginning of the twentieth century, had not yielded convincing results. In fact, all observations were consistent with an unchanging Sun, and with measurement variations induced by the Earth's atmosphere through which the measurements had been made – thus the term "solar constant" was believed to be a reasonable connotation for the flux of radiation emitted by the Sun (e.g. Smith and Gottlieb 1973). We had to wait for measurements from space before variations in the solar constant could be measured reliably.

THE STUDY OF SOLAR OSCILLATIONS PRIOR TO 1975

In the summers of 1960 and 1961, Robert Leighton (1961) and his two students Robert Noyes and George Simon made observations that fundamentally altered the

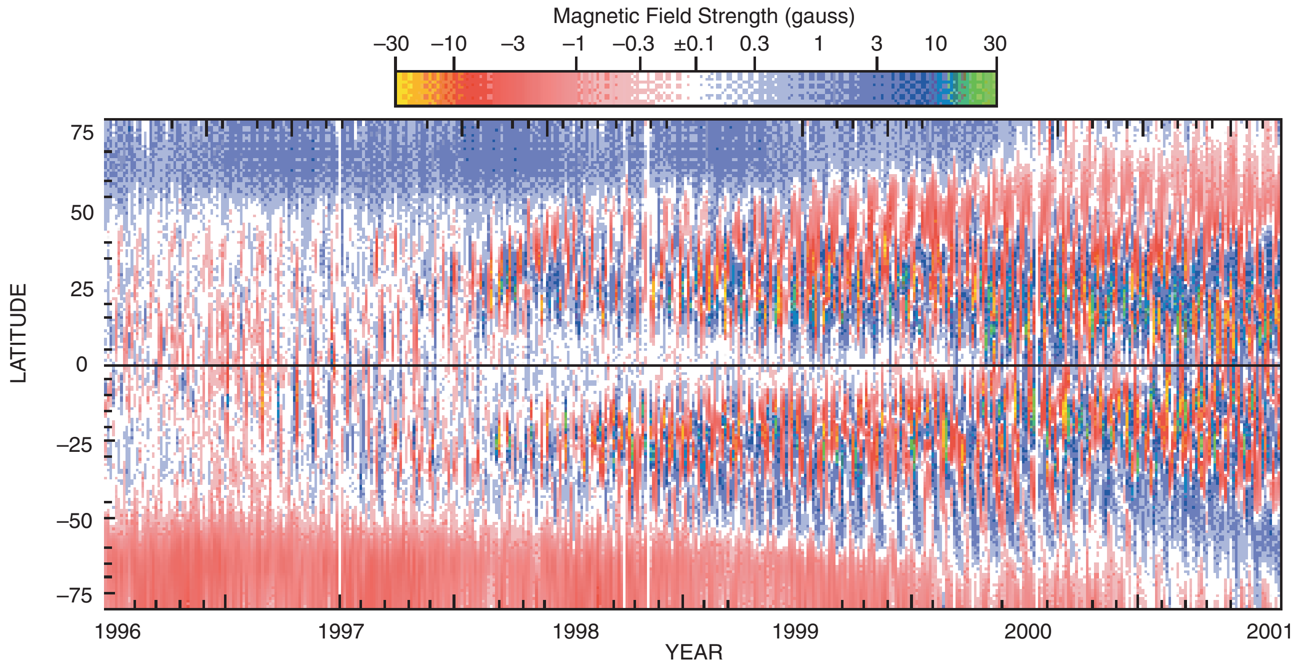

Figure 2 Solar magnetic flux during the beginning of the present activity cycle. The strong fields in darker colors can be seen to migrate toward the equator while the weaker fields share the structure of the large-scale polarity pattern which organizes the corona. Each Carrington rotation is evident in the vertical banded structure, which is a symptom of the variation in field strength (and activity) with longitude. (Courtesy of Dave Hathaway.)

study of solar physics (Leighton *et al*. 1962). By optical subtraction of spectroheliograms obtained in the wings of a spectrum line they discovered both supergranulation (Simon and Leighton 1964) and the 5-minute oscillations (Noyes and Leighton 1963). Although we still do not really understand supergranulation in detail, we do know that it structures and rearranges magnetic fields and plays a crucial role in the outer atmosphere. The discovery of the 5-minute oscillations spawned helioseismology and the modern study of the solar interior.

The spatial or temporal coverages of the early studies were limited, masking the essential nature of the oscillations. The true nature of the oscillations as the superposition of millions of trapped acoustic waves was at first neither understood nor exploited.

Frazier (1968) claimed that 5-minute solar oscillations were trapped acoustic modes. He suggested to Ulrich that he carry out a calculation which resulted in the understanding that the observations were of evanescent waves that are vestiges of waves propagating in the interior of the Sun and being reflected beneath the observing layer (Ulrich 1970). At about the same time Leibacher and Stein (1971) reached the same conclusion. This understanding was not fully accepted until Deubner (1975) observed the predicted ridges in power after making an observation with greater spatial and temporal resolution. These observations were soon followed by more detailed observations by Rhodes *et al*. (1977).

At about the same time H.A. Hill announced having seen oscillations in the solar diameter measurements which were interpreted as being global modes of oscillation (e.g. Hill and Caudell, 1979). The diagnostic potential was recognized immediately (Christensen-Dalsgaard and Gough 1976), and helioseismology was born. Hill *et al*. (1991) prepared a detailed review of these early observations.

HELIOSEISMOLOGY IN A NUTSHELL

The Sun is a ball of almost fully ionized gas. Its radius is about 700 Mm. The nuclear reactions that heat the Sun occur mainly in the inner 150 Mm. Up to about 500 Mm from the center, the energy is carried outward by radiation. An average photon takes about 30,000 years to traverse the radiative zone; the thermal cooling time is 1,000 times longer. The outer 200 Mm is unstable to convection, and there the energy is efficiently carried by convection up to the thin layer at the surface called the photosphere from where the energy is radiated into space. It takes several months for energy to get through the convection zone and about 8 more minutes to get to the Earth. Figure 3a is a schematic view of the main features of the Sun.

A number of types of waves can propagate in the Sun. Acoustic waves can propagate throughout the interior. The sound travel time through the Sun is about 2 hours. Acoustic

http://www.springer.com/978-0-7923-7196-0

The Century of Space Science
Bleeker, J.A.; Geiss, J.; Huber, M. (Eds.)
2001, XLIX, 1846 p., Hardcover
ISBN: 978-0-7923-7196-0