Global hydrodynamics of the sun

A. S. Monin

P. P. Shirshov Institute of Oceanography, Academy of Sciences of the USSR
Usp. Fiz. Nauk **132**, 123–167 (September 1980)

CONTENTS

> *"There's nothing simpler than a star."*
> *Sir Arthur Eddington.*

INTRODUCTION

Of particular interest in the field of geophysical hydrodynamics (i.e., the hydrodynamics of the natural flow of rotating baroclinic stratified fluids), which emerged after World War II, are "global" problems, which involve the analysis of hydrodynamic processes with scale dimensions of the order of an entire planet or star. Some global problems concerning the earth, for example, are the problems of the general circulation of the atmosphere, the circulation of the world ocean, the shaping of the climate, and the generation of the geomagnetic field (the magnetohydrodynamic processes which operate in the liquid shell around the earth's core).

For the sun, the principal global hydrodynamic problems are the differential rotation and the generation of the 11-yr cycle in the solar activity. These problems are evidently related, and they have attracted theoretical interest for a long time. Now that man has begun to venture forth into the solar system, an understanding of the origin and nature of the planetary and solar magnetic fields has come to be considered a necessity. Approaches for solving these problems have been pointed out, and they hold the promise of rewarding effort with success in the very near future. This situation has motivated the present review.

Interest in the differential rotation of the sun is not limited to solar physicists: A differential rotation also occurs on Jupiter and Saturn, and there are some analogous effects (the so-called jet streams) in the earth's atmosphere and oceans (the equatorial counter-currents). Many leading meteorologists, beginning with Wilhelm Bjerknes, one of the founders of modern dynamic meteorology, studied the differential rotation of the sun. Rossby offered an explanation for the equatorial acceleration of the solar rotation on the basis of a meridional transport of angular momentum against the gradient of the average zonal velocity (a negative viscosity). As Starr (1968) showed, an energy transfer from smaller to larger scales is a general characteristic of the photosphere as well as terrestrial atmospheric and oceanic flows. Golitsyn (1972, 1973) demonstrated that the problem of the general atmospheric circulation of planets was related to that of the sun by applying to the sun his scaling theory for the circulation of planetary atmospheres. The extension to the sun required some modification of this theory, since the motions in the atmospheres of the terrestrial planets arise from their nonuniform heating by the sun, while for the sun itself the driving force for the circulation is not a heating by some external source but the solar rotation and the convection caused by heating from within. Golitsyn's theory led, in particular, to the estimate that the depth of the solar convection zone was 0.2–0.3 of the solar radius, in agreement with several other estimates.

The differential rotation of the sun is of course accompanied by the expenditure of energy to overcome viscous forces (most important is the turbulent viscosity caused by the small-scale convective motions in the granules and supergranules); without some mechanism to replenish this energy, the angular rotation velocities at the various heliographic latitudes would become equal in a few rotations of the sun, according to estimates. According to current ideas, this replenishment mechanism is the meridional and radial transport of angular momentum in the convection zone by giant convection cells, which affect the rotation of the sun. These cells form a spiral macroturbulence (in which the curl of the velocity is not orthogonal to the velocity itself). Proof of this hypothesis should come from numerical simulations of the overall circulation of the convection zone, with an individualized description of the giant convection cells and of the parametrized turbulent viscosity; these calculations would be analogous to the numerical simulations which have been carried out for the general circulation of the earth's atmosphere, with an individualized description of cyclones and anticyclones.

Turning to the 11-yr cycle, we note that its obvious consequences in the earth's magnetosphere and its possible consequences in many other phenomena on the earth have attracted a great deal of interest (perhaps unwarranted in some cases) from workers in a wide variety of fields. In addition to the familiar and undisputed consequences of the solar activity on the earth—the aurorae, magnetic storms, the bombardment of space vehicles and astronauts—there are some less familiar effects, such as the variations in the level of

radioactive carbon in the atmosphere (in particular, the "de Vries fluctuation" in the 16th through 19th centuries), which raise the possibility of reconstructing the history of solar activity in the past centuries and even millenia. Some investigators are seeking correlations with climatic variations.

I take the position, however, that any direct effect of the solar activity on the earth's weather has yet to be proved. This opinion is based, in particular, on the failure to find evidence of the dominant periods in the solar activity (11 yr and 22 yr) in the oscillation spectrum of the air temperature, that of the amount of precipitation, or those of several other meteorological characteristics and climatic indicators [see, for example, p. 21 in the book by Monin (1969)]. This opinion has been defended elsewhere by Khromov (1973), who analyzed in detail many pieces of evidence which seemed to link the terrestrial weather with the solar activity. Finally, Pittock (1978) has recently reviewed the statistical relationships between the solar activity (the 11-yr and longer cycles) with variations in the weather, concluding that all the existing evidence for such relationships with periods in the range 11–22 yr is based on either an incorrect use of the methods of mathematical statistics or a biased selection of data, which results in an apparent correlation between several weather effects and the solar activity. Pittock believes that, "...if in the future more data and better analysis enable the detection of statistically significant relationships, these will account for so little of the total variance in the climatic record as to be of little practical value."

We know that the 11-yr sunspot cycle is also the cycle for the polarity reversal of the solar magnetic fields, so that the generation and oscillation of the solar activity and of the solar magnetic field result from the same mechanism, and an explanation for the solar cycle should emerge from the theory for the solar magnetic dynamo. It is presently believed that the poloidal solar magnetic field forms from an existing toroidal field as the result of a spiral macroturbulence which consists of giant convection cells (and which continuously sustains the differential rotation), and a new toroidal field forms from the poloidal field because it is stretched out by the differential rotation (the energy loss which occurs in the course of these events is so large that the oscillation of the magnetic field acquires a relaxation nature). Proof for these arguments, and a quantitative explanation for the differential rotation, should emerge from numerical simulations on the magnetohydrodynamics of the spiral macroturbulence in the convection zone.

These global hydrodynamic problems of the sun have not yet been solved, but significant progress has been made, by the approaches mentioned above. We will be reviewing the latest work on these problems in this paper.

1. THE SUN AS A WHOLE

For a summary of what is known about the sun the reader is referred to the collection edited by Kuiper

(1953), the book by de Jager (1959), and the paper by Pikel'ner (1966).

a) Dimensions

We see the sun in the sky as a small disk with an angular size of only half a degree (more precisely, $31'36''$ in January, $31'28''$ in July, and $1919''.26$ on the average). At the mean distance to the sun, $a_\odot = 1.4953 \times 10^8$ km (it takes light 494 sec to travel this distance), we have $1'' = 725$ km, so that the radius of the sun is $R_\odot = 695\,300$ km, or more than 100 times that of the earth. The mass of the sun is a third of a million times that of the earth: $m_\odot = 332\,958\,m_\oplus = 1.9901 \times 10^{33}$ g. Hence we find the very low mean density $\rho_{mean} = 1.408$ g/cm^3, and the gravitational acceleration at the surface of the sun turns out to be $g = Gm_\odot / R_\odot^2 = 273.98$ m/s^2—a huge figure, nearly 30 times that for the earth.

b) Rotation

The motion of visible details on the disk (primarily the dark spots) shows that the sun is rotating about its axis; the solar equator is inclined at an angle of $26°.4$ with respect to the earth's equator and at a small angle of $7°.2$ with respect to the plane of the earth's orbit (the ecliptic); and the longitude at the ecliptic of the ascending node of the solar equator is $73°.7$ (these figures refer to the epoch 1850 and change slowly over time). The direction of the rotation is the same as that in which most of the planets, including the earth, rotate around their own axes and revolve around the sun. The mean sidereal (stellar) period of the sun's rotation is 27 days (mean solar days); the corresponding angular rotation velocity is $\omega = 2.7 \times 10^{-6}$ s^{-1}, and the corresponding linear velocity at the equator is about 2 km/s (this is slow; linear surface velocities at the equator have been found to be tens of times higher for many more massive stars of spectral classes O, B, A, and F. As far back as 1863, Richard Carrington showed that the rotation of the sun was differential: There is an "equatorial acceleration," i.e., an increase in the angular rotation velocity toward the equator. An empirical expression for the variation of the rotation velocity ω with the heliocentric latitude was offered by Newton and Nunn in 1951:

$$\omega = 14°.38 - 2°.77 \sin^2 \varphi \ (\text{day})^{-1}. \tag{1.1}$$

The rotation period turns out to be about 25 days at the equator and 29 days at a latitude of $60°$. Section 3 below is devoted specifically to the rotation of the sun.

c) Radiation

From the measured value of the solar constant, $I_0 = 1.96$ cal/(cm$^2 \cdot$ min), we find the luminosity of the sun, i.e., the total power of its electromagnetic radiation, to be 3.8×10^{33} erg/s. This is equivalent to the explosion of 10^{11} megatons of TNT per second. The mean output of electromagnetic energy per unit mass of solar matter is then $\varepsilon \approx 1.93$ erg/(g$\cdot$s). The corpuscular emission of the sun (mainly electrons and protons) is much weaker, and the same is true of the neutrino emission, according to the information available. The effective radiation temperature T_*, found from the equation $\sigma T_*^4 = q_0$, where $q_0 = 6.31 \times 10^{10}$ erg/(cm$^2 \cdot$ s) is

the radiation flux density at the solar surface, is 5798 K. The sun is then a star of the yellow-dwarf class (spectral class G2). Spectral analysis shows that in the outer part of the sun, which is what is visible to us (we will call it the *solar atmosphere*), hydrogen atoms make up about 91% of the total number of atoms, helium accounts for about 9%, and heavier elements account for less than 0.1%. Correspondingly, the specific concentration of hydrogen is $X \approx 0.70$, that of helium is $Y \approx 0.27$, and that of heavier atoms is $Z < 0.025$ [see Aller (1961)].

d) Atmosphere

To describe the solar atmosphere we will make use of a reference level at the spherical surface $r = R_\odot$, which corresponds to the sharp visible limb of the solar disk. This level is that at which the optical depth is $\tau = \tau_{5000Å} = 0.003$ (i.e., the intensity of radiation at a wavelength $\lambda = 5000$ Å emitted in the radial direction from this level would be reduced by absorption in the atmosphere above this level by a factor of $e^{0.003}$). The brightly emitting gas below this level from which the radiation still manages to pass through the atmosphere above without suffering severe absorption is called the *photosphere*. It is about 400 km thick. Above the photosphere is the *chromosphere*, which can be seen during eclipses as a "grassfire": a dark red strip with teeth (*spicules*) at the top. It appears a few seconds before the sun emerges from behind the lunar limb. The chromosphere is of the order of 15000 km thick. Above the chromosphere is the bright *corona*, which can be seen during total eclipses. Its rays may extend out to a distance of several $R_\odot$.

1. Photosphere. Theoretical models of the photosphere are based on the assumption that the photosphere is in a state of *local thermodynamic equilibrium*. In other words, each infinitesimal volume of the photosphere is assumed to emit and absorb radiation as if it were an absolute black body at the same temperature. In particular, therefore, Kirchhoff's law holds: The ratio of the spectral absorptivity and the spectral emissivity is independent of the nature of the optically active substances and is a universal function of the wavelength and the temperature (the Planck function). Under this assumption, the various models for the photosphere are found through a joint integration of the hydrostatic equations and the radiation-transport equations, with some particular dependence of the absorptivity on the temperature, the pressure, and the wavelength. The principles underlying the construction of these models are set forth in the book by Mustel' (1960) and by M. Minnaert's paper in Kuiper's collection (1953).

From the models emerge the following conclusions about the photosphere:

1. At depths greater than 350–400 km, the gas becomes essentially opaque.

2. The photosphere is an extremely low-density gas, with a typical density of $(1-3) \times 10^{-7}$ g/cm^3 and a typical pressure of the order of 5–150 mbar.

3. Slightly above the photosphere, the temperature goes through a minimum of about 4170 K, at an altitude

of 200 km; this is the coldest part of the sun. The velocity of sound, c, is 7–8 km/s at the temperatures prevailing in the photosphere. The photosphere is hydrostatically stable not only above this cold layer, where the temperature increases with altitude, but also below it, down to a depth of about 350 km. Further into the interior, the temperature gradient becomes "superadiabatic," and this region is the upper part of the subphotospheric convection layer. The local scale height, $H = p/\rho g = RT/\mu g$, in the photosphere is 100–150 km ($\mu \approx 0.6$ is the average molecular weight); the adiabatic temperature gradient $\gamma_a = g/c_p$ is 17 deg/km. The Väisäläi-Brunt period $2\pi/N$, where

$$N^2 = \frac{g}{T}\left(\gamma_a + \frac{\partial T}{\partial z}\right),$$

falls off with altitude in the photosphere; its typical values are 200–210 s, and it reaches a minimum (of about 180 s) at $200 < z < 500$ km.

The electron gas pressure p_e falls off with altitude in the photosphere (the only increase is between 300 and 600 km), but the ratio p_e/p, which is a measure of the degree of ionization of the medium, behaves differently: It is at a minimum (below 1×10^{-4}) over most of the photosphere (at $200 \text{ km} > z > -200$ km, below which the ionization increases rapidly with depth), while above the photosphere it increases with altitude, reaching $p_e/p \approx 1/3$ at 15 000 km.

The normal (unperturbed) photosphere has a grainy structure (*granulation*). The granules are bright polygonal areas, separated by darker lines, with diameters ranging from 200 to 1300 km (the average is 760 km), with a brightness 10–30% greater than the average background, and with an average lifetime of 8–10 min. The granules have conventionally been interpreted as reflecting *Bénard cells*, i.e., laminar-convection cells which develop in the upper part of the subphotospheric convection layer and which penetrate by virtue of their inertia to some altitude in the hydrostatically stable photosphere. The gas rises at the centers of these cells and drops along their edges, at velocities of the order of 0.3 km/s [according to the Doppler shift of spectral lines; the same measurement technique has revealed oscillations in the solar atmosphere with a period of about 300 s (i.e., of the order of the Väisäläi-Brunt period in the region above the photosphere), which evidently reflect internal waves]. The intergranular network, on the other hand, is apparently the structure of the magnetic field.

One of the most important discoveries in recent years was the discovery by Severny *et al.* (1976) of an oscillation in the solar atmosphere with a period of 160 min. They found this oscillation by analyzing voluminous data on the velocities along the line of sight in the solar atmosphere at the equator and at the poles. Brooker *et al.* (1976) found the same period independently from the Doppler shifts of Na and Ca absorption lines; they also found periods of 29, 40, and 58 min. These periods do not correspond to natural modes, according to calculations based on existing models for solar structure [see, for example, Vorontsov and Zharkov (1978)] so that, possibly, these models will need to be revised.

2. Chromosphere. Let us take a brief look at the situation higher in the solar atmosphere. The chromosphere can be seen not only at the limb of the solar disk but also over the entire disk, in some particular spectral line which is emitted only, or predominantly, by the chromosphere. Examples are the red Balmer line of hydrogen, Hα 6563 Å, and the lines K3934 Å and H 3968 Å of ionized calcium, Ca$^+$II. The chromosphere is described as consisting of a lower part, in which the ionization of hydrogen is still slight, and an upper part, which is highly ionized [according to H. C. Van de Hulst, the corresponding altitude interval is $z = 7000$–14 000 km, with a temperature increase from 2.5×10^4 to 3×10^5 K and a decrease in the number density of electrons from 5×10^9 to 5×10^8 cm^{-3}; see Kuiper (1953)]. In the upper chromosphere we can no longer expect a local thermodynamic equilibrium; the kinetic temperature of the electron gas, for example, is not equal to the ionization temperature or to the excitation temperature of the spectral lines.

The chromosphere is apparently very inhomogeneous, with a filamentary structure; intense turbulence has been detected in it, with velocities ranging from 5 km/s at an altitude of 500 km to 20 km/s at 5000 km.

Spicules and *supergranulation* are constantly observed in the unperturbed chromosphere. The spicule thickness is of the order of 500–600 km; they reach an average altitude of 7500 km at the equator and 7800 km at the poles; their lifetimes are 2–5 min; and the gas is rising in them with a velocity of the order of 20 km/s. They are found in the greatest number at the poles (where the number is 30% higher than at the equator) and in the smallest number at a latitude of 35° (where the number is 10% lower than at the equator). The spicules are straight and usually inclined; the polar spicules are inclined toward the equator, while in the active regions at latitudes of 20°–40° the spicules are inclined toward the nearest pole, probably following the direction of the magnetic field.

The supergranulation is a large-cell structure in the solar atmosphere, which can be seen both in the photosphere and (particularly well) in the chromosphere. In the photosphere the correlation functions of horizontal brightness inhomogeneities in white light have a secondary maximum at a scale dimension of 12 000 km, which corresponds to supergranules, in addition to the one corresponding to the granules. In the chromosphere it is possible to observe in, for example, Ca$^+$ lines, a chromospheric grid formed by chains of large "nodules" (bright calcium flocculi), which are groups of small nodules (granules) and which are distinguished by a descent of the gas at about 1–1.5 km/s. The supergranules are of the order of 30 000 km in diameter, so that there are only about 2500 of them on the solar disk. Their lifetime is about a day. They are interpreted as manifestations of large convection cells in the interior.

There are apparently also some even larger, *giant*, convection cells (detected by their magnetic field; see the end of Section 2, where it is stated that regions of opposite magnetic-field polarity alternate over longitude with a predominant wave number $m = 6$). These giant

cells should be sensitive to the solar rotation, so that they may be a manifestation of Rossby waves in the outer convection zone. The supergranules and giant cells are morphologically similar to the clusters of convection clouds in the tropical atmosphere of the earth above the oceans. Finally, it is suspected that there may exist very long-lived *supergiant cells*, which are responsible for the so-called "active longitudes" in the spot-formation activity on the sun and for the sector structure of the solar wind (see Subsections 2a and 3b).

3. *Corona*. With a coronograph, the solar corona can be observed on any clear day; it is not necessary to wait for an eclipse. Out to a distance $r \sim 2R_\odot$ it is possible to see light emitted by the corona in the lines of highly ionized iron, nickel, and calcium (the most intense lines are the green line Fe XIV 5302.86 Å and the Fe XIII ultraviolet lines). The corona also emits x rays and radio waves (especially at meter wavelengths). Out to $r \sim 4R_\odot$, it is possible to see white sunlight scattered and polarized by free electrons in the corona; out even further, we see the inner zodiacal light (light scattered by interplanetary dust). The light from the inner corona is 10^6 times fainter than that from the disk, and the light from the outer corona is 10^3 times fainter yet. The corona consists of a very highly ionized plasma, far from local thermodynamic equilibrium.

In the inner corona $[r = (1.03-1.2)R_\odot]$ the temperature rises from 1×10^6 to 1.5×10^6 K; the electron number density N_e drops from 2×10^8 to 4×10^7 cm^{-3}; and the mass density $\rho \sim m_H N_e$ (where $m_H = 1.67 \times 10^{-24}$ g is the mass of the proton) drops from 5×10^{-16} to 1×10^{-16} g/cm^3. Here the scale height H is of the order of 10^3 km, and the velocity of sound is of the order of 15 km/s. At higher points, the gas temperature is constant, on the average, because of the high electron thermal conductivity. In the central part of the corona, at $r \sim 2R_\odot$, we have $N_e \sim 2 \times 10^6$ cm^{-3} and $\rho \sim 5 \times 10^{-18}$ g/cm^3; in the outer corona, at $r \sim 3R_\odot$, these values are lower by another order of magnitude (at $r \sim 10R_\odot$, we have $N_e \sim 10^4$; $N_e \sim 10^2$ at $r \sim 50R_\odot$; and $N_e \sim 2.5$ cm^{-3} at $r \sim 215R_\odot$). The corona is believed to be heated by the energy dissipated by acoustic (shock) and MHD waves, which bring about 10^{-5} of the solar radiation, and the radiative cooling is believed to be slight because of the negligible mass density.

e) Interior

There is no directly determined information of any sort on the part of the sun below the photosphere, and the structure here must accordingly be calculated theoretically, using hydrostatic equations and the equations for radiative equilibrium:

$$\frac{dp}{dr} = -\rho g, \quad g = \frac{Gm}{r^2}, \quad \frac{dm}{dr} = 4\pi r^2 \rho, \quad (1.2)$$

$$4\pi r^2 \frac{d\sigma T^4}{dr} = -\frac{3\varkappa\rho}{c^2} L, \quad \frac{dL}{dr} = 4\pi r^2 \rho \varepsilon, \quad (1.3)$$

where $\varkappa$ is the *opacity* (a weighted average over wavelength of the spectral absorptivity $\varkappa_\lambda$), and ε is the specific rate at which energy is released in thermonuclear fusion reactions. Both these quantities must be specified or calculated separately. In addition, use is made of the equation of state of an ideal gas, and the convection

TABLE I. Internal structure of the sun, according to Sears (1964).

r, 10¹¹ cm	$\frac{m}{m_\odot}$	ρ, g/cm³	T, 10⁶ K	L, 10³³ erg/s	ϰ	X
0.00	0.00	158	15.7	0.00	1.09	0.36
0.06	0.05	103	13.8	1.30	1.32	0.52
0.10	0.2	59	11.3	3.09	1.78	0.65
0.15	0.4	31.5	9.0	3.77	2.42	0.69
0.20	0.6	15.2	7.1	3.90	3.2	0.70
0.26	0.8	5.0	5.1	3.90	4.5	0.71
0.32	0.9	1.84	3.9	3.90	6.0	0.71
0.38	0.95	0.74	3.0	3.90	7.4	0.71
0.48	0.99	0.117	1.73	3.90	9.6	0.71
0.62	0.99955	0.0063	0.66	3.90	—	0.71

zones are treated as adiabatic [see the papers by B. Strömgren in Kuiper's collection (1953) and in the collection of Aller and McLaughlin (1965)]. The evolution of the chemical composition of a star which results from the fusion reactions can be taken into account. A calculation of this sort has been reported by Sears (1964), who found the structure described in Table I for a star of age of 4.5×10^9 yr with the mass, radius, and luminosity of the sun and with a chemical composition described by specific concentrations of $X = 0.708$ for hydrogen, $Y = 0.272$ for helium, and $Z = 0.020$ for heavy elements. According to this model, the pressure at the center of the sun is $p_c = 3.5 \times 10^5$ Mbar, so that $p_{c\odot}/p_{c\oplus} \approx m_\odot/3m_\oplus$. At $r \lesssim 0.7R_\odot$ in the interior, there is a state of radiative equilibrium with a hydrostatistically stable stratification (according to certain solar models, there is a small convective core with a radius of about $0.05R_\odot$). The temperature at the center of the sun according to this model is 15×10^6 K. Curve I in Fig. 1 shows the subadiabatic temperature profile in the zone of radiative heat transfer (where there is no convection, but there may be internal waves). Curve II shows the superadiabatic profile $T(r)$ in the upper convection zone; curve III shows the subadiabatic profile $T(r)$ in the zone of radiative heat transfer in the photosphere; curve IV shows the profile in the chromosphere, where the temperature increases with altitude because of the energy absorbed from mechanical waves (the upper part of regions IV and V corresponds to the boundaries of supergranules); curve V shows the profile in the corona, where heat conduction also comes into play; and curve VI shows the profile in the outer corona, where the temperature falls off with altitude because of the heat carried off by the solar wind and because of heat conduction.

Figure 2 is a full logarithmic plot of the radial profile of the mass density. In the interior, the logarithm of the density varies comparatively slowly: from $\rho \sim 150$ g/cm^{-3} at the center of the sun to a value 1000 times
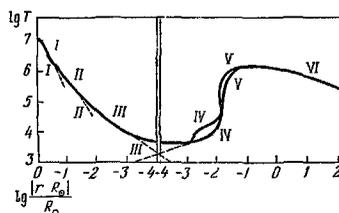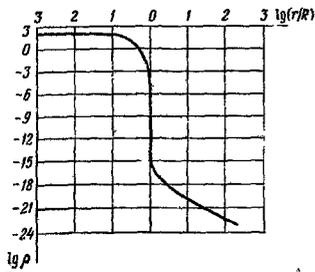

FIG. 1. Radial profile of the temperature in the sun.

FIG. 2. Radial profile of the density in the sun.

lower at the lower boundary of the convection zone $[r \sim (2/3)R_\odot]$. In the convection zone, the density falls off by another factor of 10^6, reaching a level of the order of 10^{-7} g/cm^{-3} at the lower boundary of the photosphere. The density falls off almost discontinuously (by a factor of 10^6 over a distance of 2000 km) in the photosphere and lower chromosphere, near the visible surface of the sun. In the corona and beyond, in the solar-wind region, on the other hand, the density decrease is much slower.

1. Solar neutrinos. At present the temperature at the center of the sun may be lower than that calculated according to the model in Fig. 1. In this connection, let us review the results of some neutrino experiments by Raymond Davis in 1967–1968 and 1972, in which an attempt was made to measure the flux density of electron neutrinos, ν_e, formed in the solar interior in the course of fusion reactions. Davis made use of an idea which had been proposed in 1946 by B. M. Pontecorvo: to use the "inverse beta decay" of chlorine, $Cl^{37} + \nu_e \rightarrow Ar^{37} + e^-$, which is sensitive to the highest-energy particles in the flux of solar neutrinos. (Their energies are above 0.814 MeV. These high-energy neutrinos can be formed in one of the various possible branches of a proton fusion reaction, in which a He^3 nucleus combines with a He^4 nucleus to form Be^7; the latter combines with a proton to form B^8; and the beta decay of this boron nucleus, $B^8 \rightarrow Be^8 + e + \nu_e$, generates the neutrinos. The efficiency of this neutrino production is an extremely strong function of the temperature—it is proportional to T^{20}—so that the number of these neutrinos in the solar radiation would give us a reliable value for the temperature in the solar interior.) The detector used by Davis consisted of 610 metric tons of liquid CCl$_4$, which was placed in a deep mineshaft for shielding against cosmic rays. The radioactive argon, Ar^{37}, formed in the CCl$_4$ was accumulated for 100 days and then extracted by blowing helium through the volume and adsorbing the argon on activated charcoal at $-196°$C. The number of decaying Ar^{37} atoms was measured with a proportional counter.

According to the theory based on the model which has the temperature at the center of the sun equal to 15×10^6 K, the counter in this experiment should have found 45 Ar^{37} atoms per run. The actual result, however, was negative: No more than eight Ar^{37} atoms were detected. This number corresponds to the estimated background noise level. Attempts can be made to explain this negative result in several ways: hidden systematic errors in the experiment; errors in the labora-

tory data on the rates of neutrino reactions, which were used in the calculations; changes in the properties of the solar neutrinos during their passage to the earth (for example, Pontecorvo has suggested that there may be "oscillations," consisting of the conversion of electron neutrinos into muon neutrinos or into antineutrinos); or, finally, an incorrect solar model (the temperature in the interior might, for example, be lower than that predicted by the models). This latter possibility was pursued by W. A. Fowler and, later, by D. Ezer and A. G. W. Cameron and by Sakurai. According to their hypotheses, there is a periodic mixing of the solar interior (the period would be 10^8 yr) which results in a state with a reduced temperature (over a time of the order of 10^7 yr) in which there are decreases in both the neutrino emission (by an order of magnitude) and the photon luminosity (by 20–35%). This is the state which would be prevailing today. However, I believe that it would be at least premature to adopt this particular hypothesis in order to explain the negative results of Davis's neutrino experiment.

2. Convection zone. The outer part of the sun, starting at about $0.3R_\odot$ (or even $0.4R_\odot$, according to some estimates), turns out to be convective,[1] since the adiabatic index $\gamma = (d\ln p/d\ln\rho)_a$ there is approximately unity (during compression of the plasma, energy is expended primarily on ionizing hydrogen, rather than raising the temperature), and the absorptivity of the plasma is high. Thus the dimensionless temperature gradient in the medium, $\nabla = \partial\ln T/\partial\ln p$, is higher than the adiabatic gradient $\nabla_a = (\gamma - 1)/\gamma$, and convection sets in. A more detailed calculation of the structure of the convection zone can be carried out by the following procedure: 1) Equate the specific dissipation of the kinetic energy of the convective motion, $\sim w^3/l$, to $(gq/c_p \rho T)Q$ [l is the mixing length; w is the velocity of the convective motion; and $q = c_p \rho w T'$ is the convective heat flux, about 6×10^{10} erg/(cm$^2 \cdot$s), with a correction $Q = 1 - (\partial\ln\mu/\partial\ln T)$, where μ is the average molecular weight of the solar plasma, $T' = l(T/H)(\nabla - \nabla_1)$ is the heating of the convection cells, $\nabla = (H/T)\partial T/\partial z$ is the normalized temperature gradient in the medium, and ∇_1 is that in the rising convection cell]. 2) Require that the sum of the convective heat flux, q, and the radiative heat flux, $(16.3)(\sigma T^4/\varkappa\rho H)\nabla$, be constant (equal to σT_*^4). 3) Equate the radiative heat loss of a spherical convection cell of radius l over its lifetime l/w, which is $(16/3)(\sigma T^4/\varkappa\rho H)(\nabla - \nabla_1) \cdot 4\pi l^2 l/w$, to the change in its enthalpy, $(4.3)\pi l^3 c_p \rho l(T/H)(\nabla_1 - \nabla_a)$. The resulting three equations are used, with fixed values of T, ρ, and l, to find ∇, ∇_1, and w. Then the hydrostatic equation is used to calculate the stratification of the convection

---

[1] In the outer layers of the sun, the lithium level is very low, and the beryllium level is slightly low, in comparison with their abundances in other parts of the solar system. The situation may be a consequence of the consumption of lithium and beryllium in thermonuclear reactions in the solar interior. According to calculations, lithium is consumed at distances $r < 0.63R_\odot$ from the center of the sun, while beryllium is consumed at $r < 0.47R_\odot$. It can thus be concluded that at $r > 0.63R_\odot$ the radial mixing is important, while at $0.63R_\odot < r < 0.47R_\odot$ it is not.
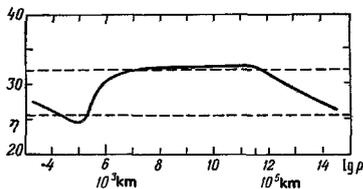
FIG. 3. Vertical profile of the dimensionless specific entropy in the convection zone below the photosphere, from Böhm (1967).

zone, i.e., the depth profiles of T, p, and ρ in this zone. Böhm (1967) has carried out such calculations, making two hypotheses regarding the mixing length, $l = H$ and $l = |z|$, finding very nearly the same results in the two cases. The results show that the upper part of the convection zone has some "mesostratification." This stratification can be seen clearly on the profile of the dimensionless entropy in Fig. 3, which shows that over most of the convection zone the specific entropy falls off only very slightly with altitude, but in a thin upper layer about 400 km thick (with a scale height H of the order of 200–300 km) it drops sharply; it is in this least stable layer that the convection cells which induce the granulation care formed. The specific entropy increases in the photosphere, but the convection cells entering it from below continue to float up actively, until they reach an altitude where the specific entropy of the gas is higher than in these cells (inertia may carry the cells slightly above this point).

2. PHENOMENA IN THE SOLAR ATMOSPHERE

In addition to the granules, supergranules, and spicules, which are all present in the normal (unperturbed) solar atmosphere, there are also several "perturbations," i.e., inhomogeneities with relatively short lifetimes. The most important perturbations are spots, faculae, chromospheric flares, prominences, coronal streamers and holes.

a) Sunspots

These are the most apparent inhomogeneities on the surface. In fact, they can sometimes be seen with the naked eye, through smoky glass or simply smoke. Such observations date back more than 2000 yr. The invention of the telescope in 1611 was followed immediately by drawings of sunspots and the first information on their shape and abundance. Fabricius inferred the rotation of the sun from the motion of the spots over the solar disk, and Galileo found the first estimate of the rotation velocity.

The spots are generated as small dark *pores* with a diameter of 2–4 seconds of arc. A typical mature spot consists of a dark central *umbra*, with an average diameter of about 17 500 km and a brightness 20–30% of that of the surrounding unperturbed photosphere, enclosed by a less dark, annular *penumbra*, with an average outer diameter of about 37 000 km and a brightness 75–80% of the background brightness. The area of a spot is typically of the order of 10^{-4} of the area of the visible surface. Balancing this reduced emission

level in the umbra and penumbra is a ring of elevated brightness (about 3% above the background level) around the spot, at an average distance of 50 000 km from the center of the spot. This ring is defined most sharply in the upper chromosphere above the spot.

An individual sunspot lasts for a time ranging from a few days to a few months. High-resolution photographs of the spots at the center of the visible disk reveal an important difference in the granulation at the surface of the photosphere outside the spots and in their umbras and penumbras. The granules in a penumbra are thin, bright, radial filaments with a diameter on the order of 300 km, which end abruptly at the boundary of the umbra and which have lifetimes ranging from half an hour to several hours (i.e., several times the lifetime of the granules of the unperturbed photosphere). On photographs taken at especially high resolution the granules are usually seen as bright points even in the umbra; they are slightly smaller than the granules of the unperturbed photosphere and have longer lifetimes—comparable to that of the granules of the penumbra.

As early as 1769, Alexander Wilson noted that as a spot approached the western limb the eastern half of the penumbra would gradually contract and then disappear completely; when the spot reappeared at the eastern limb, the western half of the penumbra would initially be missing altogether and would then appear and gradually expand. This effect means that the optical density of the gas directly above the umbra is much lower than at the same level in the surrounding atmosphere; i.e., the surfaces of constant optical density funnel downward above a spot. Sunspots are shallow funnels at the surface of the photosphere.

In 1907, E. Walter Maunder discovered an east–west asymmetry in the distribution of spots visible on the disk: The number of spots which are observed, which form, and which emerge from behind the solar limb on the eastern side is higher than the number of spots which are observed, form, and disappear behind the limb on the western side. This effect can be attributed to an inclination of the vertical axes of the spots toward the west (for young spots, 1 day old on the average, the inclination is 0°.44; for spots seen a second time, i.e., more than 27 days old, the inclination is 7°.6), which occurs because *the angular rotation velocity of the sun increases with altitude in the photosphere and the lower chromosphere*.

The spots are dark only in contrast with the unperturbed photosphere; their apparent blackness means only that their emission in the visible range is less than that of the unperturbed photosphere. If we assume that the spots, like the photosphere in general, radiate as heated blackbodies, we can compare their radiation temperatures (which will be equal to the ordinary gas temperatures, since the photosphere is in local thermodynamic equilibrium). The average temperature in the umbra of the larger spots, according to measurements of the integrated emission intensity, is about 4200 K, or 1600 K lower than the temperature of the unperturbed photosphere.

The temperature of an umbra falls off noticeably with increasing spot area. The radiation temperatures found for a given spot in different spectral lines are different. Even for spots of the same area, however, the radiation temperatures measured in the same spectral line may vary by more than 1000 K (because, for example, the generation depths for the given line are different). The lowest umbral temperatures, below 1000 K, have been detected from the intensity ratio of the emission lines of thermally unstable TiO molecules.

Measurements of the gas velocities in spots on the basis of the Doppler shift of the centers of gravity of certain spectral lines allowed J. Evershed to conclude, already in 1909, that at the level of the unperturbed photosphere and deeper the gas flows out of a spot with a certain anticyclonic twisting, along directions defined by the filaments of the penumbra (and with an ascending component), at velocities of the order of 2 km/s. These velocities increase with depth in the photosphere to 5–6 km/s. Later measurements showed that this motion can be seen at altitudes up to 500–1000 km above the surface of the sun; above this level, the motion reversed, and at altitudes of the order of 2000 km in the chromosphere there is a cyclonic twisting flow of gas into a spot, with a descending component.

Sunspots are observed only in certain latitude intervals (in the "royal latitudes"), from 5° to 52° in each hemisphere. The overwhelming majority of spots appear between 8° and 30° (the record high latitude of 52° was noted on 13 August 1953). As early as 1889, A. Wolfer noted that an increased number of spots could be observed for several years, or perhaps even tens of years, in certain active longitude intervals. These intervals change slowly over time.

The overwhelming majority of spots appear and spend their lives as part of groups. The groups exist for an average of 10 days, but 37% of the groups have lifetimes greater than 10 days, 0.4% greater than 50 days, 0.03% greater than 100 days, and 0.01% greater than 150 days. The lifetime of a group is most sensitive to the area occupied by the group, varying roughly in proportion to this area: The lifetime of a group occupying 1×10^{-4} of the surface of the visible hemisphere is 10 days, while that of a group occupying 4×10^{-4} of the visible hemisphere is 40 days. About 3000 groups are observed over one solar cycle (which lasts an average of 11.2 years; see Section 4).

b) Magnetic fields of spots

The discovery of strong magnetic fields in spots is crucial to an understanding of the nature of spots. The magnetic fields in the solar atmosphere are measured by exploiting the Zeeman effect: the splitting of spectral lines in a magnetic field.

The splitting of certain lines in the spectra of sunspots had actually already been noted in 1866, by Sir Joseph Lockyer, but it was not until 1910 that George Hale observed opposite circular polarization of the lines in these doublets and found that they were produced by a normal longitudinal Zeeman effect. In this manner,

Hale discovered the strong magnetic fields in sunspots. Hale's measurements at the Mt. Wilson observatory over the years 1908–1924 (the results were published in collaboration with S. B. Nicholson in 1938) laid the foundation for what we know today about the magnetic fields of sunspots. Hale used the iron line at 6173.348 Å almost exclusively, and the error of his measurements was ± 50 G. His students Horace W. Babcock and Harold D. Babcock later constructed a more accurate magnetograph, which used a spectrograph with two narrow slits, which were aligned at the wings of the spectral line under study (the iron line at 5250.218 Å in most cases), where there are inflection points. This device detected the difference between the photocurrents corresponding to the light transmitted through these slits and modulated by a circular-polarization analyzer.

The accuracy of this magnetograph was of the order of 0.1 G; the horizontal resolution was 70″, later improved to 23″.

It was learned that *all sunspots have strong magnetic fields.* The magnetic lines of force at the generation depth of the spectral lines which undergo the Zeeman splitting run roughly perpendicular to the solar surface at the center of a spot and bend outward with increasing distance r from the center, so that the inclination angle θ varies approximately as described by $\theta = 90°(r/r_0)$, where r_0 is the radius of the outer edge of the penumbra. The lines of force within the penumbra follow the filamentary granules and become approximately horizontal at the outer edge of the penumbra.

The magnetic field at the center of a spot, H_c, increases with increasing spot area, from something of the order of 100 G for the smallest spots to perhaps 400 G for the large spots. The field H falls off with distance from the center of the spot approximately in accordance with $H = H_c [1 - (r^2/r_0^2)]$. The magnetic flux through the surface of a typical spot is of the order of 10^{21} Mx (1 Mx = 1 G·cm²). As a spot grows, the magnetic flux in its increases rapidly at first and then remains approximately constant for much of the spot lifetime; the area of the spot increases, and the field decreases.

Differences in the nature of the magnetic field in spot groups lead to the classification of groups as "unipolar," "bipolar" (or "dipolar"), or "multipolar." The overwhelming majority are bipolar; for example, 91% of the 6384 groups observed over the years 1919–1946 were bipolar, 8.6% unipolar, and only 0.4% multipolar.

Clearly, we should focus our attention on the bipolar regions. Each has a "preceding" spot (in the sense of the direction of the solar rotation), which we will designate as "p," and one or more "following" spots ("f"). The magnetic polarities of the p and f spots in a given bipolar region are always opposite (and the magnetic fluxes in the p and f parts of a bipolar region are approximately equal).

On the average, the p spots are larger than the f spots and have longer lifetimes; the magnetic flux in the p spots is three times that in f spots, on the average. The bipolar groups are usually oval, with the long axis fp inclined with respect to the circles of

601   Sov. Phys. Usp. 23(9), Sept. 1980

A. S. Monin   601

latitude in such a manner that the p spot is closer to the equator than the f spot. The actual inclination angle decreases with increasing age of the bipolar group and with decreasing latitude. At latitudes in the range 5°–9°, the average angle is 3°; in the range 10°–14° it is 5°.5; in the range 15°–19° it increases to 6°.5; in the range 20°–24° it increases to 9°.4; in the range 25°–29° it increases to 11°.9; and finally, in the range 30°–34° it is 14°.9. The spots within a bipolar group move with respect to each other. During the first 5–7 days in the life of a bipolar group, the p spot moves westward rapidly (the average velocity is of the order of 0°.2 per day) but in a decelerating manner; the f spot, on the other hand, remains at rest or moves slowly eastward. The group becomes longer, frequently reaching a length of 10°–12°. Then the p spot stops and sometimes moves slowly eastward, frequently returning to its original position; at this time the f spot disappears. Finally, the group as a whole changes latitude slightly (of the order of 1° per solar rotation), moving toward the equator if the group is in the latitude band $|\varphi| < 16°$ and moving toward the poles otherwise.

The formation of a bipolar group may possibly result from the ascent of some part of a magnetic force tube (which has a small inclination with respect to a circle of latitude in the horizontal plane), from below the granulation layer under the photosphere. The ascent may result from a local intensification of the magnetic field (because of, for example, a nonuniform stretching of the tube by the differential rotation). With the natural tendency toward an equalization of the total pressure, $p + H^2/8\pi$ (the hydrodynamic pressure plus the magnetic pressure), the result would be a decrease in the pressure p because of a displacement of some of the gas from the field-intensification region. A further result would be a decrease in the gas pressure, which would lead to an Archimedes buoyant force. An attempt can be made to explain the differences in the shape and evolution of the p and f spots in terms of the effects exerted on the p and f regions of the magnetic force tube by the angular rotation velocity of the photospheric gas, which increases with altitude (as mentioned earlier in connection with the east–west asymmetry of the number of visible spots). This mechanism for the formation of bipolar groups is described in more detail by, for example, Babcock (1961).

The spots form in bipolar groups probably as a result of a modification and partial suppression of the convective motion in the granulation layer below the photosphere caused by a strong magnetic field (which makes it exceedingly difficult for plasma to move across magnetic lines of force): The reduction of the convective heat transfer results in a cooling of the gas in the photosphere and thus the appearance of relatively cool spots.

c) Faculae

These are comparatively long-lived bright regions near sunspots (small faculae may exist without being associated with spots, but *no spots exist without faculae*), which can be seen both at the limb of the disk in white light (i.e., in the photosphere) and over the entire disk in the bright chromospheric lines, especially $H\alpha$ and Ca II. (Measurements are usually made in the K_{232} part of the K Ca II line, at whose center there is a narrow absorption line K_3, which is enclosed by a narrow emission band K_2 and an even broader, deep absorption band K_1. The bright elements of chromospheric faculae are called *flocculi*.)

Photospheric faculae have a granular structure; their brightness averages 10% higher than that of the unperturbed photosphere, and the difference becomes 40–45% in the brightest facula granules and sometimes as much as 150%. Granules in faculae live longer (about an hour, on the average) than those in the unperturbed photosphere. These granules, with dimensions of the order of 1000 km, are grouped in cells 4000–6000 km in size and form chains 5000–10 000 km wide and of the order of 50 000 km long. These chains continue the chromospheric network into the photosphere.

Most faculae appear in a latitude interval which includes the "royal latitudes" important for sunspots but extends 10° further toward the poles and 5° further toward the equator. In addition, a few polar faculae are observed, distributed roughly uniformly over the latitude band $|\varphi| > 68°$ (these are small, rounded, bright spots, 1800–3000 km in size, with an average lifetime of half a day; they are apparently associated with the polar coronal streamers and appear most frequently at minimum solar activity).

The large faculae associated with sunspots frequently appear a few hours, or even a few days, before the spots themselves and continue to be observed long after the spots have disappeared (this is especially true of the p and f spots in bipolar groups), so that their lifetimes average three times those of the spots. These large faculae frequently remain visible over several solar rotations.

Chromospheric faculae are present in all bipolar spot groups with magnetic fields stronger than 2 G. They have long lifetimes, reaching 200–300 days. They frequently form elongated regions whose western ends are nearer the equator than the eastern ends. The angles made by these elongated regions with the circles of latitude are noticeably larger than those in the case of the fp axes of bipolar groups. The temperature in these faculae increases with altitude more rapidly than that in the unperturbed chromosphere, so that their temperature contrast increases with altitude. In the $L\alpha$ 1215 Å line, this contrast is twice that in the K_3 Ca II line, and in x radiation (i.e., in the corona) the faculae are 70 times brighter than the background.

d) Chromospheric flares

These events are sudden, short-lived increases in the emission intensity of regions in the chromosphere near sunspots; they can be seen best in the $H\alpha$ line and in the lines of Ca$^+$. The area of a flare is $(1-12) \times 10^{-4}$ of the area of the visible hemisphere (1.6×10^{-4}, on the average). The relatively common small flares are rounded, while the relatively rare large ones are elongated formations, with dimensions of the order of

602   Sov. Phys. Usp. 23(9), Sept. 1980

A. S. Monin   602

$10^4 \times 10^5$ km and a filamentary structure. The integrated brightness may increase by a factor as high as ten during a flare. Flares last from a few minutes to a few hours (the average lifetime is 20^m), and this time tends to increase with increasing flare area, although there is a large scatter in this dependence. The brightness increases rapidly; the maximum brightness is maintained only briefly (less than 1^m); and the brightness decays comparatively slowly (the decay time averages three times the rise time).

Flares are not rare events: One flare occurs, on the average, during each 7^h of the lifetime of a spot group (flares occur particularly frequently between the 8th and 15th days in the life of a group), so that some 30–50 flares occur as a spot group travels across the disk. At maximum solar activity, a flare occurs evey 2 h, and about 300 occur per rotation. Flares usually appear at faculae and begin with a rise of the bright elements of faculae.

Two-thirds of the flares are flat, extended, *static* formations, but one-third exhibit intense motion, with velocities of hundreds of kilometers per second (the parabolic velocity at the sun is 618 km/s). This motion sometimes results in *surges* of matter at an arbitrary angle from the vertical, to altitudes of the order of 10^5 km; then the matter returns *along the same path*. Flares are also accompanied by x-ray emission, UV emission (primarily in the $L\alpha$ line), *large bursts* of radio emission, and the emission of corpuscular fluxes and cosmic rays.

The emission of particles implies that some mechanisms are operating to accelerate particles during flares, and these mechanisms evidently involve varying magnetic fields. Direct measurements of the magnetic fields have shown that the flares seem to be associated with zones between magnetic fields of different polarity, in which zones the magnetic field changes most rapidly. Finally, flares are *condensations* of chromospheric plasma, with a density 10^2–10^4 times that of the surrounding chromosphere, and this point is crucial to an understanding of their nature. This condensation also explains the sharp increase in emission (and the subsequent cooling of the flare region). The picture which emerges is thus one in which the flares form as the result of a compression of gas in the region in which increasing magnetic fields "collide."

e) Prominences

These are ribbon-shaped, cool condensations of gas in the inner corona, which appear as bright outcrops at the solar limb but appear as dark filaments on the disk, in the light of the cores of Fraunhofer lines (e.g., the hydrogen Balmer line Hα). A typical quiescent prominence is of the order of 200 000 km long (a length as great as 1 900 000 km has been observed), 50 000 km high, and no more than 6000 km wide, consisting of filaments of the order of 1000 km in diameter. The average lifetime of a prominence is three solar rotations. They are observed, first, in the latitude band 10°–40°, which includes the sunspot band but extends further toward the poles, and, second, at high latitudes (in a polar-prominence crown).

The prominences associated with spots frequently are shaped like arcs, arches, fountains, or loops, probably following magnetic lines of force. The material condensed in them is moving predominantly downward. Prominences almost always arise near a facula, and usually on the polar side of the facula; the filament is usually directed toward the spot (80% of the filaments are directed toward a p spot), and at the time the spot forms the filament is apparently aligned with a meridian. Then the filament stretches out (by an average of 100 000 km per rotation) and acquires a characteristic shape, a meridional arc deformed by the differential rotation; the high-latitude part of the filament approaches an east–west direction. Furthermore, the filament moves slightly poleward (at a velocity ranging from 2°.3 of latitude per rotation in the latitude band 0°–10° to 0°.8 of latitude per rotation in the band 51°–69°). This displacement occurs through the erosion of the low–latitude end of the filament and the growth of the high-latitude end. Finally, nearly all the filaments go through a stage of sudden disappearance: In the course of a few hours, they move downward into the chromosphere or upward into the corona, or they simply contract. A few days later, they reappear in their previous form.

At the altitude of the prominences, 50 000 km, the temperature in the corona is of the order of 700 000 K, and the electron number density is of the order of 2 $\times 10^8$. In the prominences, at the same pressure, the temperature will be approximately inversely proportional to the density: The typical temperature is 15 000 K, and the typical electron density is 2×10^{10}. The mass density is 2×10^{-14} g/cm^3, and at a velocity of the order of 5 km/s the kinetic-energy density is 0.002 erg/cm^3. A magnetic field with this energy density would have an intensity of 0.2 G, while a field of intensity of the order of 3 G will exist at an altitude of 50 000 km above a spot with a field of 3500 G. The motion of matter in the prominences is thus controlled by the magnetic field.

The increase in the density of the cool matter in the corona may result from a thermal instability; the emission there is proportional to the square of the density. For free–free transitions of protons and electrons, the emission is proportional to $T^{-1/2}$, while for free–bound transitions it is proportional to $T^{-3/2}$, so that the density increase and the temperature decrease should continue as long as the "heat of condensation" can be radiated off.

f) Coronal streamers and holes

Streamers with a base thickness of the order of 7000 km and with an electron density five times that in the surrounding medium can be distinguished in the corona. The streamers are closely associated with prominences (they frequently coincide at the solar limb) and apparently are usually coronal shells around filaments. The latter determine the general isophot (or isolux) distribution in the white corona. Among the common structural forms in the corona are *fan streamers*: rays which are inclined above a system of concentric arcs which envelop a relatively dark dome around a pro-

tuberance.

An important event in solar science in the mid-1970s was the discovery of radial regions of reduced density in the corona—coronal holes—and their association with fast streams of solar wind [see the review by Zirker (1977)]. Most of what we know about holes was learned from the Skylab observations, although ground-based observations have also contributed. These holes form in the photosphere but are controlled by the magnetic field of the inner corona. Although they tend to group in the polar regions of the sun, their rotation is rigidly tied to the rotation near the equator. The divergence of the magnetic lines of force in the corona acts as a Laval nozzle: The motion of the gas in the divergence regions is accelerated to a supersonic level, and the gas density decreases.

g) Centers of activity

These are the regions on the sun in which the events discussed above occur, in a certain order.

Small facula spots appear on the *first* day and rapidly stretch out in the east–west direction, with their western ends closer than their eastern ends to the equator. The first spot appears in the western part of the facula region on the *second* day. A second spot forms in the eastern part of the facula region on the fifth day; a multitude of small spots, the first chromospheric flares, surges, and short-lived filaments appear between the p and f spots. On the 11th day, the spots reach their maximum development and are surrounded by large penumbras; the facula region continues to grow; and the flare activity increases. By the 27th day, all the spots except the p spot have disappeared, and the flares have become less frequent, but the facula region continues to grow, and the filament on its polar side, which makes an angle of about 40° with the meridian and is pointed toward the p spot, stabilizes.

By the 54th day, no spots remain; the facula region becomes less bright and is bisected by a filament, which reaches a length of about 100 000 km and is approaching an east–west orientation. By the 81st day, the facula converts into a widely spaced network, while the filament is still being lengthened by the differential rotation. By the 108th day, the facula has dissipated, leaving only the chromospheric fine structure and the filament, which has reached its maximum length. By the 135th day, all that is left is a shortened filament, displaced slightly toward a pole. Between the 162nd

and 270th days, the filament reaches the polar prominence crown and merges with it. The guiding factor in the evolution of a center of activity seems to be the *magnetic field*.

h) Solar magnetic field

An overall description of the solar magnetic field was given by Babcock (1981). The results of global mappings of the solar magnetic field were published by Bumba and Howard (1965); see also the review by Howard (1967). Starr and Gilman (1965–1968) have offered a magnetohydrodynamic interpretation of these results. Figure 4 shows an example of a global map of the magnetic field (recorded in August 1959, right after a solar maximum). This map reveals many bipolar and unipolar spot groups with typical magnetic fields 1–20 G (outside spots). In the latitude band $|\varphi| < 40°$, the bipolar groups are predominant. At each fixed latitude in this band, regions of positive and negative polarity alternate, with a predominant longitudinal wave number $m = 6$. The axes of these regions are inclined with respect to circles of latitude, with their western ends closer to the equator than their eastern ends. Since the motion is roughly along magnetic lines of force, the contours of these regions apparently describe Rossby waves in the east–west flow. In each of the polar regions, $|\varphi| > 40°$, unipolar spot groups of the corresponding polarity are predominant, and this observation can be taken as evidence for the existence of a *poloidal* component in the magnetic field.

3. ROTATION OF THE SUN

The data on the rotation have been reviewed by Gilman (1974), Howard (1975), and Durney (1976).

a) Observational data

There are two ways to measure the rotation of the solar atmosphere: by studying the motion of long-lived inhomogeneities in this atmosphere (i.e., by studying the motion of sunspots, filaments, faculae, magnetic fields, coronal streamers, etc., although these features may have a motion of their own with respect to the rotating atmosphere) and by studying the Doppler shift of spectral lines emitted from some part of the atmosphere (although the useful signal will have to be studied against an intense background caused, in particular, by motion in the granules and especially supergranules and by scattered light from other parts of the solar at-
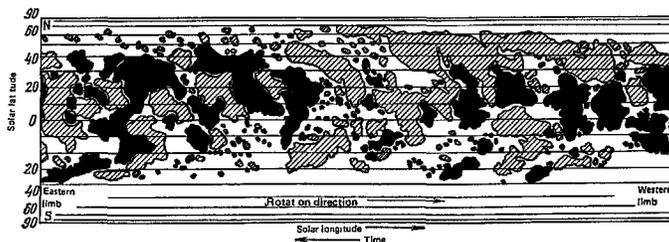
FIG. 4. Solar magnetic field over one solar rotation period (August 1959), according to Bumba and Howard. Black and shaded regions correspond to magnetic field component along the line of sight greater than 2G; black regions correspond to positive and shaded regions to negative polarity of the magnetic field.

mosphere).

The first measurements of the differential (over latitude) rotation of the visible surface (the photosphere) were carried out by following the motion of sunspots; the spots chosen for this study were long-lived spots which crossed the central solar meridian at least twice (the p spots in certain centers of activity). This method can be used, of course, only with the royal latitudes. The empirical formula of Newton and Nunn [Eq. (1.1)] was constructed in this manner. The most extensive measurements of the motion of spots have been carried out on the basis of the data in the Greenwich sunspot catalog over a period of 76 yr by Ward (1964, 1965a, 1965b, 1966a, 1966b, 1966c). Ward found that, averaged over all spot groups, the rotation velocity at all latitudes is 1% higher than predicted by Eq. (1.1); the small groups are moving faster than the large groups, and those stretched out along the longitude are moving faster than the rounded groups. Ward also observed significant deviations of the motion of the spots from a uniform differential rotation, and we will discuss these deviations below. Here we wish to recall that the east-west asymmetry in the visibility of spots which was noted by Maunder as well as the differences in shape and evolution of the p and f spots can be interpreted as resulting from an increase with altitude of the angular rotation velocity in the photosphere and lower chromosphere.

According to results found by Wilcox and Howard (1970), the large-scale photospheric magnetic fields of the type in Fig. 4 rotate at the same velocity near the equator as at higher latitudes—faster than the sunspots (their rotation velocity undergoes large fluctuations; the differential rotation is much less apparent in the very long-lived magnetic fields in the latitude band 10°–25° than in the short-lived fields or absent altogether). Among the photospheric formations, we might also recall the polar faculae at latitudes $|\varphi| > 60°$, which exhibit a definitely differential, and especially slow, rotation.

Turning our attention to higher layers in the atmosphere, we note that for the filaments in the inner corona a differential rotation has been detected (this rotation stretches the filaments out) which is just slightly faster than that of sunspots. Almost exactly the same rotation velocities have been found for the inner corona by spectroscopic techniques by Hansen *et al.* (1969) and, later, by several other investigators. We should emphasize here that in most cases no significant variation with altitude has been observed in the angular rotation velocity of the chromosphere and the inner corona (out to a distance of about $2R_\odot$).

Results slightly different from all those described above were obtained from the Doppler shift of spectral lines emitted from the photosphere. The most extensive Doppler measurements were carried out with the solar magnetograph at the Mt. Wilson Observatory; see Howard and Harvey (1970) and Howard (1971). These measurements were carried out daily, beginning in 1966, at an array of points covering the entire solar disk. The error in the measurement of the velocity

component of the photospheric gas along the line of sight was 10 m/s. These measurements revealed, first, a slower rotation of the photosphere at all latitudes than that found from the motion of sunspots (the rotation periods were about a day longer). Second, significant variation with time was observed in the rotation velocity; the changes reached 10–20% of the average value. Some of these variations apparently result from the superposition of large-scale motions on the mean differential rotation; these large-scale motions are radial, at 50–75 m/s, with horizontal scale dimensions much larger than supergranules, and horizontal, at 40–50 m/s, with a longitudinal period of the order of 25°. Third, it was found that there is a slight tendency toward an increase in the rotation velocity with decreasing solar activity.

All these results are summarized in Fig. 5, from which we see that the rotational velocities of the photospheric magnetic fields, sunspots, filaments, and coronal streamers are approximately the same, at least in the equatorial region; these velocities apparently characterize primarily the rotation of the magnetic lines of force. The lines of force probably are rotating along with the highly ionized matter below the granules and supergranulation and in turn entrain the highly ionized, low-density gas of the upper chromosphere and corona. In the photosphere, on the other hand, and in the lower chromosphere, the gas is only weakly ionized, and it almost completely escapes entrainment by the magnetic lines of force moving through it. On the average, it lags behind the motion of the lines of force, as can be seen from the Doppler shift of the photospheric emission lines. If we accept this picture, the primary effect is a differential rotation of the matter with the magnetic field below the layers of granulation and supergranulation; the rotation of the photosphere and the lower chromosphere is a secondary effect, resulting from the downward transfer of momentum from the originally rotating layers provided by convective motion of the supergranules and granules. Also a secondary effect is the rotation of the upper chromosphere and the corona, caused by the rotation of the magnetic lines of force, perhaps with some effect of a downward momentum transfer by acoustic, gravitational, and MHD waves.
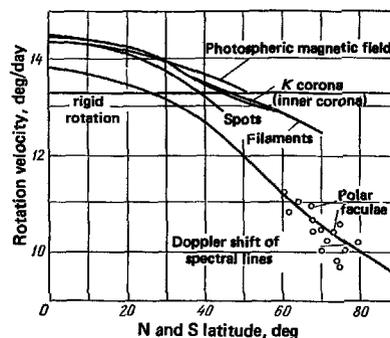

FIG. 5. Angular rotation velocity of the solar atmosphere at various latitudes.

b) Rotation of the solar wind

The solar plasma propagates along with the magnetic field far beyond the corona, although in an extremely low-density form, as the "solar wind" [see the review by Lüst (1972)]. The density is so low at a few $R_\odot$ from the sun that there are essentially no collisions between the particles making up this plasma. The rotation of the solar wind around the sun can be inferred from the directions in which cometary tails point [see Brandt and Heise (1970)] and from direct measurements from space vehicles. Measurements carried out in 1962–1967 near the earth's orbit ($215R_\odot$ from the sun) on the Mariner 2 probe and on several satellites (Vela 2, 3, and 4 on IMP 1) showed that the average particle density in the solar wind at this distance is 5 cm^{-3} [of the positively charged particles, 95% are protons and 4.5% are α particles; the charge flux density carried by the positively charged particles is 1.75×10^8 electron charges/(cm$^2\cdot$s); the average proton temperature is 4×10^4 K, and this temperature is twice as high along the lines of force of the interplanetary magnetic field as across them; the α-particle temperature is four times the proton temperature; and the electron temperature is three times the proton temperature]. According to the data acquired on the Vela satellites, the average velocity of the solar wind is 320 km/s, and this velocity makes an angle of 1.5° with the radius vector from the sun, in the direction of the solar rotation. This deviation corresponds to an azimuthal velocity of 10 km/s, five times that at the solar surface (the same value is found from the directions of cometary tails). From angular momentum conservation, $rv_\lambda = $ const, we find the azimuthal velocity v_λ at the earth's orbit to be $r/R_\odot = 215$ times smaller than at the solar surface. Consequently, there must be a significant angular momentum flux F moving away from the sun. Let us estimate this flux.

The average magnetic field in the solar wind near the earth's orbit turns out to be about 5×10^{-5} G. The energy density of the magnetic field is much smaller than the kinetic-energy density, so that the field is carried off by the solar wind, and its lines of force are Archimedes spirals coming from the sun. If the radial velocity of the solar wind, v_r, is constant, then this field can be described by Parker's equations (1958):

$$B_r = B_{r0}\left(\frac{r_\odot}{r}\right)^2, \quad B_\theta = 0, \quad B_\lambda = B_r\frac{\omega r \sin\theta}{v_r}, \qquad (3.1)$$

where B_{r0} is the field on the same magnetic line of force at the surface of the sun, and ω is the angular velocity of rotation of this line of force. This field creates a retarding torque (i.e., the outgoing angular momentum flux) which acts on the sun:

$$F = r^3 \int_0^\pi \sin^2\theta\, d\theta \int_0^{2\pi} \frac{B_r B_\lambda}{4\pi}\, d\lambda \approx \frac{4}{3}\frac{\omega}{v_r}(r_\odot^2 B_{r0})^2, \qquad (3.2)$$

where the right side is calculated under the assumption $B_{r0} = $ const. With $v_r = 400$ km/s and $B_{r0} = 2$ G (this figure leads to $B_r = 5 \times 10^{-5}$ G at the earth's orbit) we find $F = 4 \times 10^{30}$ dyn·cm. Furthermore, since there is azimuthal motion with a velocity v_λ in the solar wind, there is an additional transport of angular momentum by gas (and by the waves propagating through the gas, primarily MHD waves). Consequently, F can reach (7–8)

$\times 10^{30}$ dyn·cm. If the sun is rotating approximately as a rigid body, its resultant angular momentum will be of the order of 2×10^{48} g·cm^2/s; one-tenth of this angular momentum will correspond to the outer convection layer, and the torque F will be able to stop the rotation of the entire sun in 10^{10} yr and that of only the convection layer in 10^9 yr.

Schatten (1973) presents several arguments to support the assertion that over much of the solar cycle (especially at minimum activity) the solar wind is coupled by magnetic lines of force to the polar regions on the sun, so that the retarding torque F is applied to these polar regions. Then if the turbulent viscosity η in the differentially rotating layer were not too high, the torque F would be able to create a differential rotation $\delta\omega$ ($\delta\omega = 3 \times 10^{-7}$ s^{-1} is the difference between the angular rotation velocities in the equatorial and polar zones) over a time $\delta t = I\delta\omega/(F/2)$, where I is the moment of inertia of the rotating layer in the polar zone. Assigning this volume a mass $m = 0.002 m_\odot = 4 \times 10^{30}$ g, we find the moment of inertia $I = (1/2)mr_\odot^2 = 10^{52}$ g·cm^2, and with $F/2 = 5 \times 10^{30}$ dyn·cm we find $\delta t = 2 \times 10^7$ yr; i.e., the observed differential rotation would be established comparatively quickly. To reach this conclusion we have had to assume a small viscosity; specifically, the viscous stress $(\eta' r_\odot)\partial\sigma/2\theta$ must be much smaller than the stress created by the torque F, which is estimated to be 6×10^{-2} dyn/cm^2. Since we have $(1/r_\odot)\partial\sigma/\partial\theta \approx 3 \times 10^{-7}$ s^{-1} in the differential rotation, the viscosity η must be much smaller than 2×10^5 g/(cm·s). The average value of the turbulent viscosity in the convection layer of the sun, however, is estimated to be $\eta \approx 5 \times 10^9$ g/(cm·s), or six orders of magnitude larger than the value assumed by Schatten. Consequently, contrary to Schatten, we must recognize that the convection layer reacts to the retarding effect of the solar wind essentially as if it were a rigid object, so that this retardation, even though applied predominantly to the polar regions, cannot cause a differential rotation. Köhler (1970) notes that with this large viscosity the observed vertical flux of angular momentum in the convection layer can be provided by a very small vertical gradient in the angular rotation velocity. Specifically, from the condition

$$\eta r^2 \sin^2\theta\, \frac{\partial\omega}{\partial r} = \frac{F}{4\pi r_\odot^2}, \quad \frac{\partial\omega}{\partial r} \approx \frac{\Delta\omega}{0.2 r_\odot}, \qquad (3.3)$$

with $\theta = \pi/2$ and $\eta = 5 \times 10^9$ g/(cm·s) we find $\Delta\omega/\omega \sim 10^{-6}$.

In summary, although the angular momentum carried off by the solar wind is responsible for slowing the rotation of the sun over its entire evolution (and explains why the outer layers of the sun, at least, are rotating slowly at the present), this process is apparently unimportant for the differential rotation of the sun.

c) Oblateness

While the rotation of the solar wind is accessible to direct measurement, we must rely completely on indirect evidence if we wish to study the rotation deeper in the interior of the sun. One piece of evidence may be the oblateness of the sun's surface along its rotation axis, $\delta r/r_\odot$, where $\delta r = r_{eq} - r_{pol}$ is the difference be-

606   Sov. Phys. Usp. 23(9), Sept. 1980

A. S. Monin   606

tween the equatorial and polar radii. If we ignore all the magnetic and hydrodynamic stresses in the surface region, except for the stresses caused by the rotation of this region, $\omega = \omega(R)$, where $R = r\sin\theta$, we can find the oblateness from the condition for the equilibrium of the surface region:

$$\nabla p + \rho\nabla\Phi = 0;$$

$$\Phi = \Phi_0 - \int_0^R R\omega^2(R)\,dR; \tag{3.4}$$

$$\Phi_0 \approx -\frac{Gm_\odot}{r}\left[1 - J_2 P_2(\cos\theta)\left(\frac{r_\odot}{r}\right)^2\right],$$

where Φ_0 is the gravitational potential, written as the sum of a monopole and a quadrupole [J_2 is a measure of the quadrupole gravitational moment, and $P_2(x) = \frac{1}{2}(3x^2 - 1)$ is a Legendre polynomial], and Φ is the effective potential, which incorporates the centrifugal-acceleration potential. From the first equation in (3.4) we see that the surfaces of $p = \text{const}$ and $\Phi = \text{const}$ coincide. Taking the curl of this equation we find $\Delta\rho \times \Delta\Phi = 0$; hence the surfaces of $\Phi = \text{const}$ and $\rho = \text{const}$ coincide [and, since the molecular weight is constant, these surfaces also coincide with the constant-temperature surfaces, $T = T(p, \rho) = \text{const}$]. We know that the visible edge (limb) of the sun, defined by the condition that a ray tangent here will penetrate a medium with a unit optical thickness, agrees very closely with the surface $\rho = \text{const}$. The limb is thus also a surface of $\Phi = \text{const}$. Then, assuming the oblateness to be small, and assuming $\omega \approx \text{const}$ in (3.4), we find

$$\frac{\delta r}{r_\odot} \approx \frac{3}{2}J_2 + \frac{1}{2}\frac{\omega^2 r_\odot}{g}. \tag{3.5}$$

For the solar surface, the second term is 0.93×10^{-5} and corresponds to a difference between the equatorial and polar radii of only 7 km or only 0.01″ on the visible disk. Nevertheless, Dicke and Goldenberg (1967) undertook an effort to measure this tiny quantity. They covered the sun with a rotating disk with two diametrically opposite slits of slightly different size, which transmitted light (they used two different parts of the spectrum) from narrow strips along the limb (they used three different strip widths: 6.5″, 12.8″, and 19.1″ on the visible disk). Analysis of the transmitted light signals led to the value $\delta r/r = (5.0 \pm 0.7)\cdot 10^{-5}$, i.e., to a difference of 0.05″ between the equatorial and polar radii on the visible disk. (With a slit width of about 20″, the variation in the transmitted light flux was of the order of $0.05/20 = 2.5 \times 10^{-3}$, which is apparently much larger than the variation in the brightness of the sun at the limb between the equator and the poles. This latter variation is extremely small, according to measurements by Dicke and Goldenberg, corresponding to a temperature difference of no more than 3 K between the equator and the poles.)

Dicke and Goldenberg interpreted their result as the discovery that the solar gravitational field has a significant quadrupole moment, which could be caused by a rapid rotation (with a period of 1 or 2 days) and thus by significant oblateness of a dense solar core. However, more recent, and extremely careful, measurements by Hill *et al.* (1974) have failed to confirm any significant contribution of J_2 to the solar oblateness in

(3.5). Then Dicke and Goldenberg's conclusion that there is a rapidly rotating solar core, and thus a decrease in the angular rotation velocity with increasing radius, has not been confirmed. We might note that these studies provoked much interest and have generated an extensive literature.

d) Hydrodynamics of the differential rotation

Turning now to attempts to derive a theoretical explanation for the observed differential rotation of the outer layers of the sun, we first note that the maintenance of a differential rotation against viscous forces requires an angular momentum flux toward the equator, caused by both "circulation wheels" in the meridional planes and "turbulence," i.e., convection in the form of granules, supergranules, and giant convection cells and/or Rossby waves. Giant cells subject to the effects of solar rotation and Rossby waves in the solar convection zone are nearly the same thing. To calculate this angular-momentum flux we should use the hydrodynamic equations for turbulent motion. For the convection zone of the sun, the hydrodynamic equations can be written in the Boussinesq approximation. The first part of this system of equations consists of the familiar equations of motion in the spherical coordinate system r, θ, λ (r is the radius, θ is the complement of the latitude, and λ is the longitude) which is rotating at the average angular rotation velocity of the sun, $\Omega = \text{const}$. Then we write the result found by taking the average of these equations, using the following notation: v_r, v_θ, v_λ are the velocity components; $P = (\rho_*/\rho_0)(p - p_0)$ is the normalized deviation of the pressure from its static value $p_0 = p_0(r)$, which is related to the static density $\rho_0 = \rho_0(r)$ by the hydrostatic equation $\partial p_0/\partial r = -g\rho_0$, where $g \approx g_0(r_0/r)^2$ is the acceleration due to gravity (g_0 is the value at the solar surface, $r = r_0$) and ρ_* is the potential density, defined by

$$\frac{g}{\rho_*}\frac{\partial\rho_*}{\partial r} = \frac{g}{\rho_0}\left(\frac{\partial\rho_0}{\partial r} + \frac{g\rho_0}{c_0^2}\right) \equiv N^2.$$

Here $c_0 = c_0(r)$ is the static value of the velocity of sound; ζ is the normalized deviation of the specific entropy from its static value, linearized in the deviations of the pressure and the density (the normalization is carried out by dividing by the specific heat at constant pressure); and τ_{ij} is the viscous stress tensor. The second part of our system of equations is the continuity equation, which we will take in the form of the condition that the velocity field be divergence-free (the condition that the mass flux be divergence-free would be more accurate). The third part is the entropy evolution equation, in which we incorporate, among the nonadiabatic factors in the convection zone, only the heat influx due to the molecular thermal conductivity, $c_p\rho_0\chi$:

$$\frac{\partial\zeta}{\partial t} + \frac{1}{r^2}\frac{\partial}{\partial r}r^2 v_r\zeta + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta v_\theta\zeta + \frac{1}{r\sin\theta}\frac{\partial}{\partial\lambda}v_\lambda\zeta$$
$$= \frac{N^2}{g}v_r + \frac{1}{\rho_0 T_0}\,\text{div}\,\rho_0\chi\nabla T. \tag{3.6}$$

To calculate the characteristics of the turbulent motions, we should take the average of these equations, assuming that the average characteristics (which we will denote by superior bars) are independent of the time and the longitude, so that they can depend only on

r and θ. Averaging the continuity equation, we find

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\overline{v}_r+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\overline{v}_\theta=0,\qquad(3.7)$$

from which we see that the average meridional circulation can be described by means of the stream function ψ, defined by

$$\overline{v}_r=-\frac{1}{r^2\sin\theta}\frac{\partial\psi}{\partial\theta},\qquad\overline{v}_\theta=\frac{1}{r\sin\theta}\frac{\partial\psi}{\partial r}.\qquad(3.8)$$

Taking the average of the equations of motion, and ignoring the molecular momentum fluxes in comparison with the convective fluxes, we obtain Reynolds equations:

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\overline{v}_r^2+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\overline{v_r v_\theta}-\frac{\overline{v}_\theta^2+\overline{v}_\lambda^2}{r}$$
$$=\Omega^2 r\sin^2\theta+2\Omega\sin\theta\overline{v}_\lambda-\frac{1}{\rho_*}\frac{\partial\overline{p}}{\partial r}+g\overline{\zeta}$$
$$-\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2 b_{rr}+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta b_{r\theta}-\frac{b_{\theta\theta}+b_{\lambda\lambda}}{r}\right);\ (3.9)$$

$$\frac{1}{r^3}\frac{\partial}{\partial r}r^3\overline{v}_\theta\overline{v}_r+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\overline{v}_\theta^2-\frac{\operatorname{ctg}\theta\overline{v}_\lambda}{r}$$
$$=\Omega^2 r\sin^2\theta+2\Omega\cos\theta\overline{v}_\lambda-\frac{1}{\rho_*r}\frac{\partial\overline{p}}{\partial\theta}$$
$$-\left(\frac{1}{r^3}\frac{\partial}{\partial r}r^3 b_{\theta r}+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta b_{\theta\theta}-\frac{\operatorname{ctg}\theta b_{\lambda\lambda}}{r}\right);\ (3.10)$$

$$\frac{1}{r^3}\frac{\partial}{\partial r}r^3\overline{v}_\lambda\overline{v}_r+\frac{\theta}{r\sin^2\theta}\frac{\partial}{\partial\theta}\sin^2\theta\overline{v}_\lambda\overline{v}_\theta$$
$$=-2\Omega\sin\theta\overline{v}_r-2\Omega\cos\theta\overline{v}_\theta-\left(\frac{1}{r^3}\frac{\partial}{\partial r}r^3 b_{\lambda r}+\frac{1}{r\sin^2\theta}\frac{\partial}{\partial\theta}\sin^2\theta b_{\lambda\theta}\right),$$
$$(3.11)$$

where $b_{ij}=\overline{v_i'v_j'}$. Here the primes denote fluctuations, i.e., *deviations from the average values. Analogously,* averaging puts the entropy evolution equation in (3.6) in the following form:

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\overline{v}_r\zeta_1+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\overline{v}_\theta\zeta_1$$
$$=-\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2 b_{r\zeta}+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta b_{\theta\zeta}\right),\ (3.12)$$

where $\zeta_1=\overline{\zeta}-\zeta_0$, and $\zeta_0=\zeta_0(r)$ is determined from $\partial\zeta_0/\partial r=N^2/g$.

e) Reynolds stresses

Ward (1964, 1965a, 1965b, 1966a, 1966b, and 1966c) *obtained empirical estimates of the Reynolds stresses* $b_{\theta\theta}$, $b_{\lambda\lambda}$, and $b_{\theta\lambda}$ in the photosphere from data on the motion of sunspots. The values he found of the order of $\sqrt{b_{\theta\theta}}\sim 0.4$ deg/day ≈ 25 m/s and $\sqrt{b_{\lambda\lambda}}\sim 0.8$ deg/day ≈ 50 m/s (the second of these values is of the same order of magnitude as the difference between the linear velocities of the differential rotation between the equator and a latitude of 35°), while the average meridional velocity $\overline{v}_\theta$ is significantly different from zero only between the latitudes of $+5°$ and $-5°$ (where this velocity is directed toward the equator but has an average value of only ~ 0.03 deg/day ≈ 2 m/s). We see that the Reynolds stresses should be extremely important in the Reynolds equations in (3.9)–(3.11).

Ward observed a significant positive correlation between the motion of spots toward the west and their motion toward the equator, so that those spots which are moving westward more rapidly than the average have a *tendency to move toward the equator,* while those moving westward more slowly than the average have a tendency to move toward the poles or to move toward
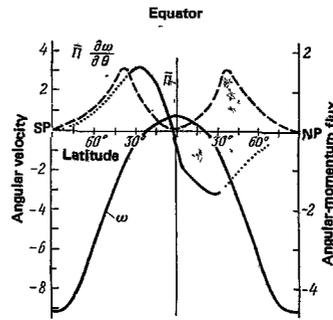
FIG. 6. Angular velocity of the differential rotation of the sun, in units of 10^{-7} rad/s; meridional turbulent angular-momentum flux $\overline{\Pi}=2\pi r^2\sin^2\theta b_{\theta\lambda}$, in units of 3.14×10^{29} g·cm²/s³; rate at which the kinetic energy of the irregular motion is converted into differential-rotation energy, $\overline{\Pi}2\omega/2\theta$, from Starr (1971).

the equator more slowly than the average (Leighton disputes the importance of this correlation, taking it to be a consequence of the specific way in which the spots behave, in particular, the inclination of the axes of spot groups along circles of latitude).

Since $r\sin\theta v_\lambda'$ *is the fluctuation in the angular momentum per unit mass with respect to the rotation axis,* $\Pi=r\sin\theta v_\theta' v_\lambda'$ is the turbulent flux of this angular momentum along the meridian (from north to south is the positive direction), and $\overline{\Pi}=\oint\Pi r\sin\theta d\lambda=2\pi r^2\sin^2\theta b_{\theta\lambda}$ is the resultant value of this flux at a circle of latitude, the correlation found by Ward ($b_{\theta\lambda}>0$ in the northern hemisphere and $b_{\theta\lambda}<0$ in the southern) means that angular momentum is transported along a meridian toward the equator, in the direction opposite to the gradient in the angular velocity of the differential rotation, $\omega=\overline{v}_\lambda/r\sin\theta$, so that the irregular motions traced out by the sunspots statistically create a negative viscosity [see Starr and Gilman (1965a, 1965b, and 1968) and Starr (1971)]. Figure 6 shows the empirical values of ω, $\overline{\Pi}$, and the rate $\overline{\Pi}\partial\omega/\partial\theta>0$ at which the kinetic energy of the irregular motions is converted into the energy of zonal flow (differential rotation) as a result of the negative viscosity, which were found by Starr (1971) from data on sunspots and from measurements of the Doppler shift.

Analogous effects are seen in the earth's atmosphere, in the latitudes around tropospheric jet streams in both the northern and southern hemispheres. By analogy with the earth's atmosphere, the correlation between
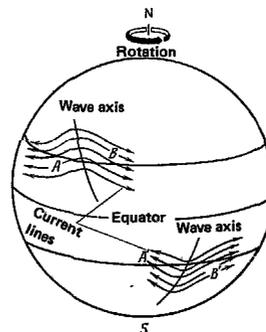
FIG. 7. Asymmetry of the waves in the east–west flow in the solar photosphere.

v'_θ and v'_λ can be interpreted as a result of an asymmetry of the waves in the east–west flow (Fig. 7): In the solar photosphere, the axes of the crests of these waves (these axes are directed toward the poles) should be inclined toward the eastern limb, so that the angular momentum and the flux of angular momentum toward the equator ahead of the waves (at points B and B') are greater than the angular momentum and its flux away from the equator behind the waves (at points A and A'). Since the electrical conductivity is extremely high in the solar plasma, the magnetic field is frozen in the moving gas, and its contours are marked by current lines, so that this inclination of the axes of the waves can also be seen in the configurations of the regions of identical magnetic-field polarity, as in Fig. 2.2. Starr and Gilman (1965a) found that in this case the Maxwell stresses $(1/4\pi)\overline{H'_\theta H'_\lambda}$ create a flux of angular momentum away from the equator toward the poles, i.e., retard the differential rotation (an irregular magnetic field is generated from the kinetic energy of the differential rotation). For the mean square values $H'_\theta, H'_\lambda \sim 7$ G, the Maxwell stresses would balance the Reynolds stresses $b_{\theta\lambda}$, but in actuality the large-scale fluctuations in the solar magnetic field seem to have a mean square value less than 7 G, and they simply reduce, but do not completely suppress, the effect of the negative viscosity.

f) Hypothesis of an anisotropic viscosity

To calculate the differential rotation, we can single out from the equations in Subsection 3d three equations for ω, φ, and ζ_1, so that Eq. (3.11) becomes

$$\frac{1}{r^3}\frac{\partial}{\partial r}r^3\left(-\frac{\omega+\Omega}{r}\frac{\partial\psi}{\partial\theta}+b_{1r}\right)+\frac{1}{r\sin^2\theta}\frac{\partial}{\partial\theta}\sin\theta\left[(\omega+\Omega)\frac{\partial\psi}{\partial r}+b_{\lambda\theta}\right]=0.$$

$$(3.13)$$

We can eliminate $\bar{p}$ from (3.9)–(3.10) by cross differentiation (assuming that the potential density ρ_* is approximately constant); we find

$$\frac{1}{r^2\sin\theta}\left[\frac{\partial(\psi,\Omega_\lambda)}{\partial(r,\theta)}+\Omega_\lambda\left(\frac{1}{r}\frac{\partial\psi}{\partial\theta}-\operatorname{ctg}\theta\frac{\partial\psi}{\partial r}\right)\right]$$
$$+r\sin^2\theta\left[\frac{1}{r}\frac{\partial(\omega+\Omega)^2}{\partial\theta}-\operatorname{ctg}\theta\frac{\partial(\omega+\Omega)^2}{\partial r}\right]$$
$$=-\frac{g}{r}\frac{\partial\zeta_1}{\partial\theta}+\frac{1}{r}\frac{\partial}{\partial\theta}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^2b_{rr}+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta b_{r\theta}-\frac{b_{\theta\theta}+b_{\lambda\lambda}}{r}\right)$$
$$-\frac{1}{r}\frac{\partial}{\partial r}\left(\frac{1}{r^2}\frac{\partial}{\partial r}r^3b_{\theta r}+\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\sin\theta b_{\theta\theta}-\operatorname{ctg}\theta b_{\lambda\lambda}\right),\quad(3.14)$$

where Ω_λ is the curl of the meridional circulation velocity,

$$\Omega_\lambda=\frac{1}{r}\left(\frac{\partial\bar{rv_\theta}}{\partial r}-\frac{\partial\bar{v}_r}{\partial\theta}\right)=\frac{1}{r\sin\theta}\left(\Delta\psi-\frac{2}{r}\frac{\partial\psi}{\partial r}-\frac{2\operatorname{ctg}\theta}{r^2}\frac{\partial\psi}{\partial\theta}\right).\quad(3.14')$$

Finally, from (3.12) we find

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\left(-\frac{\zeta_1}{r^2\sin\theta}\frac{\partial\psi}{\partial\theta}+b_{r\zeta}\right)$$
$$+\frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\left(\frac{\zeta_1}{r\sin\theta}\frac{\partial\psi}{\partial r}+b_{\theta\zeta}\right)=0.\quad(3.15)$$

The three equations in (3.13)–(3.15) contain, in addition to the three major unknowns ψ, ω, and ζ_1, some other unknowns: the Reynolds stresses b_{ij} and the turbulent entropy fluxes $b_{i\zeta}$. These equations are thus not closed. The simplest way to close them would be to express these other unknowns in terms of the major unknowns by making use of hypotheses from the semiempirical turbulence theory (although this approach introduces an

excessive arbitrariness). This method is used in attempts to explain the differential rotation in terms of the anisotropy of the turbulent viscosity caused by convection (this anisotropy results from the special role played by the vertical direction, along which the gravitational force is acting; the other direction which is singled out—the axis of solar rotation—should not be important for the motion in the granules and supergranules, for which the typical periods are very short in comparison with the period of the solar rotation). The turbulent viscosity is introduced by means of the basic hypothesis of the semiempirical theory of turbulence, according to which the Reynolds stress tensor b_{ij} is a linear function of the velocity gradients of the average motion, i.e., of the tensor $\nabla_i\bar{v}_m$. The coefficients of this linear function serve as the coefficients of the turbulent viscosity. Generally speaking, these coefficients form a fourth-rank tensor, but it is assumed in the theory of anisotropic viscosity that this tensor can be expressed in terms of a second-rank tensor K_i^j which is diagonal in spherical coordinates, with diagonal elements which are either constant or functions of r only:

$$K_\theta^\theta=K_\lambda^\lambda=sK_r^r,\quad(3.16)$$

where s is the anisotropy coefficient. In the first paper in this direction, which was unfortunately neglected by most later workers, Lebedinskii (1941) introduced Eq. (3.16) and showed that with $s\neq1$ a rotation of the sun as a rigid body could not satisfy the Reynolds equations, so that the rotation would have to be differential (since the convective turbulence results from a thermal instability, not from the energy of the average motion, and the anisotropic part of the Reynolds stress tensor, $b_i^j-(1/3)$ $\times b_\alpha^\alpha\delta_i^j$, is a linear function of $\nabla_i\bar{v}_j$, not of the rate-of-strain tensor $\Phi_{ij}=\nabla_i\bar{v}_j-\nabla_j\bar{v}_i$, so that it does not vanish when $\Phi_{ij}=0$).

This idea was introduced again by Biermann (1958, 1961) somewhat later. Reynolds equations with an anisotropic viscosity, (3.16), were written by J. Wasiutynski (1946). A detailed derivation of these equations was reported later by Elsässer (1965). Kippenhahn (1963) found an approximate analytic solution of these equations (under the boundary conditions $\bar{v}_r=b_{r\theta}=b_{r\lambda}=0$ at the upper and lower boundaries of the convection layer); Kippenhahn showed that with $s\neq1$ the solution necessarily includes a meridional circulation. With $s>1$, the gas rises at the poles, flows toward the equator in the upper part of the convection layer, and descends at the equator; in this case, the equatorial zone of the solar surface is rotating more rapidly than the polar zones, and the rotation in the convection layer accelerates with increasing altitude. In the case $s<1$, the opposite situation prevails. This research was continued by Cocke (1967), who showed that the angular rotation velocity ω turns out to be approximately a function only of $R=r\sin\theta$; Cocke found the optimum value $s\approx1.2$ for the sun. Finally, the equations were integrated numerically by Köhler (1970) with various values of s from the range 0.8–1.5 and with a viscosity $\nu\equiv K_r^r$ $=4.5\times10^{12}$ cm^2/s (and also for $s=1.2$ with viscosities from 0.1ν to 200ν; for the large viscosities, the ω distribution turned out to be approximately centrally sym-

metric, while at the low viscosities it was approximately cylindrically symmetric; the highest meridional-circulation velocity was found with $\nu = 5 \times 10^{13}$ cm^2/s, falling off at both higher and lower values of ν. With $s = 1.2$ and $\nu = 4.5 \times 10^{12}$ cm^2/s, the highest meridional-circulation velocity turned out to be less than 2 m/s, i.e., too low for detection by available equipment. We might also recall the work by Ruzmaĭkin and Vainshtein (1978), who studied the depth variation of s, finding that s falls off from the value 1.2 at the surface to 0.4 at a depth of 110 000 km and then increases to 0.6 at a depth of 190 000 km.

It is difficult to judge the plausibility of these results. For example, with values $s > 1$, which would lead to the correct sign for the differential rotation with respect to latitude at the surface, the angular rotation velocity in the convection zone would increase with altitude, but this behavior would be difficult to reconcile with the fact that the rotation of the photospheric gas lags behind the rotation of the magnetic fields, as mentioned earlier. But even if we put this objection aside, we would still have to say that the model of anisotropic viscosity, with its arbitrary elements, could hardly be taken as an *explanation* of the differential rotation. The same comment can be made about models of the type offered by Roxburgh (1970) and Durney and Roxburgh (1971) [see also Durney (1976)], which introduce an effect of the rotation on convection which causes the turbulent viscosity and the thermal conductivity to depend on the latitude (the latitude dependence is required to be of such a nature that the meridional circulation and the resulting differential rotation turn out to be similar to those observed; however, there turns out to be a large temperature difference between the equator and the poles, which has not been observed, and furthermore the stabilization of convection by rotation would have to be stronger at the equator than at the poles—and this situation would seem to be just the opposite of the probable situation).

g) Equations for the second moments

It is possible to eliminate, or at least reduce, the arbitrariness in the method used to close the Reynolds equations in (3.13)–(3.15) by dropping the attempt to somehow specify the additional unknowns b_{ij} and $b_{i\zeta}$ or to express them in terms of ψ, ω, and ζ_1 and instead by supplementing Eqs. (3.13)–(3.15) with the dynamic equations for the single-point second moments of the fluctuations b_{ij}, $b_{i\zeta}$, $b_{\zeta\zeta}$, found from the hydrodynamic equations. Although such dynamic equations would have to be simplified, so that again some arbitrariness would creep in, this arbitrariness would now have far less effect on the major unknowns, ψ, ω, and ζ_1. As one approach for simplifying the dynamic equations for the second moments, we can ignore the third moments in these equations. These simplified equations for the second moments are derived from the hydrodynamic equations, linearized with respect to fluctuations and written as follows, in tensor notation:

$$\frac{\partial V_i'}{\partial t} + \overline{V}_\alpha \nabla^\alpha V_i' + V_\alpha' \nabla^\alpha \overline{V}_i + f_i' - g\zeta' \lambda_i = -\frac{1}{\rho_*} \nabla_i P' + \frac{1}{\rho_0} \nabla^\alpha \tau_{i\alpha}', \quad (3.17)$$

$$\frac{\partial \zeta'}{\partial t} + \overline{V}_\alpha \nabla^\alpha \zeta' + V_\alpha' \nabla^\alpha \overline{\zeta}_1 = \frac{1}{\rho_0 T_0} \nabla^\alpha \rho_0 \chi \nabla_\alpha T', \quad (3.18)$$

where f' are the fluctuations in the Coriolis acceleration, and λ_i is the unit vector in the vertical direction. The physical meaning of these equations is that in the calculation of the characteristics of the average motion the interaction of this average motion with the fluctuations is taken into account, but the interaction of the fluctuations with each other is not. From (3.17) we find the following equations for the moments b_{ij}:

$$V_\alpha \nabla^\alpha b_{ij} + (b_{j\alpha} \nabla^\alpha \overline{V}_\alpha + b_{i\alpha} \nabla^\alpha \overline{V}_j) + (\overline{f_i' V_j'} + \overline{f_j' V_i'}) - g(b_{j\zeta}\lambda_i + b_{i\zeta}\lambda_j)$$

$$= -\frac{1}{\rho_*}(\overline{V_j' \nabla_i P'} + \overline{V_i' \nabla_j P'}) + \frac{1}{\rho_0}(\overline{V_j' \nabla^\alpha \tau_{i\alpha}'} + \overline{V_i' \nabla^\alpha \tau_{j\alpha}'})$$

$$\approx -c_1 \frac{b}{l}\left(b_{ij} - \frac{1}{3} b^2 g_{ij}\right) - c_2 \frac{b^3}{l}\left(\lambda_i \lambda_j - \frac{1}{3} g_{ij}\right) - 2\frac{b^3}{l}[(c_3 - 3c_4)\lambda_i\lambda_j + c_4 g_{ij}].$$

$$(3.19)$$

Here the average parts are the single-point second moments of the fluctuations, but in order to express them in terms of our additional unknowns b_{ij} we must adopt some new semiempirical hypotheses. We take the approach used in the theory of the turbulence of the boundary layer of the atmosphere [Monin (1965a, 1965b, and 1965c)]. The quantities c_1, c_2, c_3, and c_4 on the right sides are numerical constants; $b^2 = b_{rr} + b_{\theta\theta} + b_{\lambda\lambda}$ is the mean square fluctuation velocity; l is the vertical scale dimension of the turbulence; and g_{ij} is the metric tensor. The first term on the right describes the tendency of the pressure fluctuations to move the system toward isotropy; the second, in contrast, describes the tendency toward anisotropy at the boundaries of the convection zone; and the third describes the anistropic viscous dissipation of turbulence energy.

Then we find the following equations for the moments $b_{i\zeta}$ from (3.17)–(3.18):

$$\overline{V}_\alpha \nabla^\alpha b_{i\zeta} + b_{\alpha\zeta} \nabla^\alpha \overline{V}_i + b_{i\alpha} \nabla^\alpha \overline{\zeta}_1 + \overline{f_i' \zeta'} - g b_{\zeta\zeta} \lambda_i$$

$$= -\frac{1}{\rho_*} \overline{\zeta' \nabla_i P'} + \frac{1}{\rho_0} \overline{\zeta' \nabla_i^\alpha \tau_{i\alpha}'} + \frac{1}{\rho_0 T_0} \overline{V_i' \nabla^\alpha \rho_0 \chi \nabla_\alpha T'}$$

$$\approx -c_5 \frac{b}{l} b_{i\zeta} - c_6 \frac{b^2}{l} \sqrt{b_{\zeta\zeta}} \lambda_i - c_7 \frac{b^4}{l} \sqrt{b_{\zeta\zeta}} \lambda_i. \quad (3.20)$$

Finally, we find the following equation for $b_{\zeta\zeta}$ from (3.18):

$$\frac{1}{2r^3 \sin\theta} \frac{\partial(\psi, b_{\zeta\zeta})}{\partial(r, \theta)} + b_{r\zeta} \frac{\partial \zeta_1}{\partial r} + \frac{b_{\theta\zeta}}{r} \frac{\partial \zeta_1}{\partial \theta} = \frac{1}{\rho_0 T_0} \nabla^\alpha \rho_0 \chi \overline{\zeta' \nabla_\alpha T'}$$

$$-\frac{\chi}{T_0} \overline{\nabla_\alpha T' \cdot \nabla^\alpha \zeta'} \approx -c_8 \frac{b}{l} b_{\zeta\zeta}. \quad (3.21)$$

Now Eqs. (3.13)–(3.15) and (3.19)–(3.21) do constitute a closed system of equations for the unknowns ψ, ω, ζ_1, b_{ij}, $b_{i\zeta}$, $b_{\zeta\zeta}$, if we specify the vertical scale dimension of the convection turbulence, l, in some fashion. These equations contain six numerical constants, c_1, c_3, $c_4' = c_4 - (c_2/6)$, c_5, $c_6 + c_7$, and c_8. According to Monin (1965a, 1965b, and 1965c), these constants can be expressed in terms of the values of certain characteristics of the turbulence in the boundary layer of the atmosphere, in the limit of neutral stratification. They turn out to be $c_1 \approx 0.21$; $c_3 \approx 0.035$; $c_4' \approx 0.0025$; $c_5 \approx 0.2$; $c_6 + c_7 = -0.00226$; $c_8 = 0.0292$. As the boundary conditions at the upper and lower boundaries of the convection layer, $r = r_0$ and $r = r_1$, we should apparently assume a zero mechanical stress, while at the lower boundary, $r = r_1$, we should assume a given constant value $q_1 = q_0(r_0/r_1)^2$ for the vertical heat flux [$q_0 = 6.31 \times 10^{10}$ erg/(cm^2·s) is the radiation flux density at the surface of the sun]; from this heat flux we can de-

termine the vertical entropy flux $b_{r\zeta} \approx [(\varkappa - 1)/\varkappa]q/p$ ($\varkappa = c_p/c_v$ is the ratio of specific heats, and p is the pressure). The 13 equations in (3.13)–(3.15) and (3.19)–(3.21) have been integrated numerically by L. M. Simuni and the author [with $r_1 = (2/3)r_0$ and $l = \varkappa(r_0 - r)$] on a grid with six layers of identical thickness in the convection zone of the sun and ten latitude zones between the equator and a pole. The resulting angular velocity of the differential rotation at the solar surface, $\omega + \Omega$, increases from 0.964Ω at the pole to 1.143Ω at the equator; at $\theta = 0$–$7\pi/20$ it has maxima at the depth $r = r_0 - [(r_0 - r_1)/6]$, and near the equator it has maxima at the solar surface (falling off slightly with depth; at the equator for example, it falls off from 1.1426Ω at the surface to 1.1413Ω at $r = r_1$, while at the pole it falls off from 0.9641Ω below the surface to 0.8933Ω at $r = r_1$). The meridional circulation forms a single cell with surface flow from the pole toward the equator and with a flow in the opposite direction at the lower boundary of the convection zone. These results agree with results found by Kippenhahn for the case of an anisotropic viscosity with $s > 1$. Furthermore, in our case the quantity ζ, is zero at the surface, the equator, and the pole, falling off with depth, reaching a minimum at $r = r_1$ and $\theta = 3\pi/10$.

h) Numerical simulations

We have just discussed one approach for calculating the differential rotation of the sun: using the average hydrodynamic equations—the Reynolds equations in (3.13)–(3.15) with some semiempirical hypotheses or auxiliary equations for the second moments of the hydrodynamic fields. Another approach would be to integrate the unaveraged hydrodynamic equations numerically until the solutions reach a statistically steady state and then take the time average of the solutions (within the statistically steady state) and also take the average over the longitude. This approach may be called the "numerical-experiment method," and it has also been referred to as the "method of non axially-symmetric models." This approach has been developed by several workers [see the review by Gilman (1974)]; here we would like to cite the papers by Davies-Jones and Gilman (1970 and 1971) and a series of later papers by Gilman (1972, 1973, 1975, 1976, 1977, and 1978).

The early papers [Davies-Jones and Gilman (1970) and (1971)] dealt with convection in a cylindrical gap which was uniformly heated from below and which was rotating about its symmetry axis with a gravitational force acting parallel to this axis (with Taylor numbers $Ta = 4\Omega^2 d^4/\nu^2$ in the range 10^2–10^6, with $\Omega = 2.6 \times 10^{-6}$ s^{-1} as the average angular rotation velocity of the sun, with $d = 0.2R_\odot = 1.4 \times 10^{10}$ cm as the thickness of the convection zone, and with $\nu = 10^{12}$–10^{14} as the turbulent viscosity coefficient). The first and second order perturbations were calculated analytically. The numerical calculations carried out by Gilman (1972, 1973) for an equatorial cylindrical gap agreed best with the solar data with the value $Ta = 3 \times 10^4$ (here the energy of the differential rotation becomes equal to the energy of the convection cells), but these calculations led to an unrealistically strong dependence of the radiation flux

density at the solar surface on the latitude, with a maximum at the equator. The meridional circulation at the surface was found to be directed from the equator toward the poles, opposite the direction found in calculations with an anisotropic viscosity with $s > 1$. The rotation of the sun in the convection zone was found to accelerate with increasing altitude.

Gilman (1975) was the first to study the actual geometry (a spherical shell, of thickness $0.2R_\odot$ in Gilman's paper). Originally, Gilman used only the linear equations (with the boundary conditions that there were no stresses, and the temperature was constant, at the boundaries of the shell). The amplitudes of the longitudinal Fourier components of the perturbations (with wave numbers $m = 0, 1, \ldots, 24$) were calculated numerically on a grid of points in a meridional plane. Among the most unstable solutions were many modes with large values of m and with peaks near the equator, which at large values of Ta became "rollers" with axes running parallel to the rotation axis of the sun. These rollers carried momentum toward the equator. In addition, there were a few modes with small values of m and with peaks near the poles, corresponding to circumpolar chains of eddies. However, even preliminary calculations from the nonlinear equations [Gilman (1976)] showed that at Prandtl numbers $Pr = \nu/\chi$ of the order of unity the equatorial acceleration of the rotation of the solar surface occurs only with a predominance of equatorial modes [at Rayleigh numbers $Ra = g_0 \alpha d^3 \delta\theta(\nu\chi)^{-1}$ in the range $0.84\,Ta^{2/3} < Ra < 78\,Ta^{2/3}$]; this result led to the prediction that the vertical heat flux would depend significantly on the latitude, contrary to observations.

In the detailed calculations from the nonlinear equations carried out by Gilman (1977), the Rayleigh numbers were chosen so small that convection was possible, but the convection cells would be significantly affected by the rotation; here the ratio of the Coriolis frequency to the typical rate of growth of the convection cells,

$$F = 2\Omega \left(\frac{g\alpha\delta\theta}{d} \right)^{-1} = \sqrt{\frac{Pr \cdot Ta}{Ra}}, \qquad (3.22)$$

should be of the order of unity or greater (for granules, $F \sim 3 \times 10^{-4}$; for supergranules, $F \sim 10^{-1}$; for giant cells, $F \sim 1$–10). Most of the calculations were carried out for $Pr = 1$, $Ta = 10^5$, and $Ra = (1$–$4) \times 10^4$ (this is 4–16 times Ra_{cr}). With $Ra = (1$–$2) \times 10^4$, a differential rotation with the correct sign was found: There was an increase in ω toward the equator and upward (the increase in the upward direction probably occurred only below the supergranule layer, 2×10^4 km thick, in which ω increases with depth). With $Ra = 4 \times 10^4$, a differential rotation of the opposite sign was found, but in both cases the radiation flux density reached maxima at the equator and at the poles. These variations in the radiation flux density over latitude can be reduced significantly, if Ra is not too large, by replacing the boundary condition of a constant temperature at the lower boundary of the spherical shell by the condition that the radial heat flux remain constant [Gilman (1978)]. Figure 8 shows one of the resulting $\omega(r, \theta)$ distributions, this one for $Pr = 1$, $Ta = 8 \times 10^5$, $Ra = 3 \times 10^5$, and $d = \odot/3$. This distribution shows that these alternative boundary conditions lead
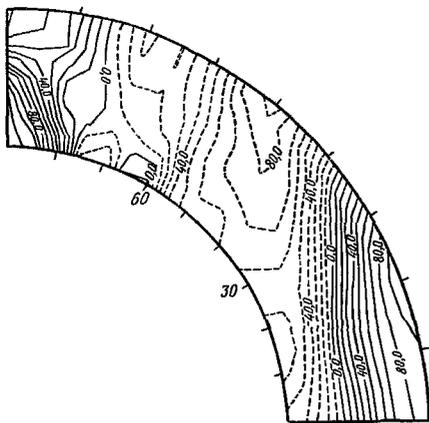
FIG. 8. The distribution $\omega(r, \theta)$ for Pr $=1$, Ta $=8\times10^5$, Ra $=3\times10^5$, and $d=R_\odot/3$. This distribution was obtained by Gilman under the boundary conditions of absence of stress at the boundaries of the spherical shell, and constancy of the temperature at the upper boundary of this shell and of the radial flux at the lower boundary.

FIG. 9. Variations in the Wolf number over the period 1610–1974.

to encouraging results. We might note that, according to Gilman's results, if the boundary conditions of zero stress are replaced by the condition that there be no slippage at the boundaries of the spherical shell (especially the upper boundary) there would be sharp decreases in the convection and in the extent to which the rotation is differential, so that the conditions prevailing in the liquid core in the interior of the earth are quite different from those in the sun, and the mechanisms for the geomagnetic and heliomagnetic dynamos may be quite different.

4. THE SOLAR CYCLE

The solar activity manifested by perturbations in the solar atmosphere (which were described in Section 2) varies with time. This is true primarily of the number of sunspots and their distribution over the solar disk. Already in 1843, Samuel Schwabe observed that the spot-formation activity on the sun varies approximately periodically, with a period of about 10 yr.

a) Wolf numbers

In 1847, the Zürich astronomer Rudolf Wolf introduced the number $W=k(N_1+10N_2)$ as an index of the solar activity; here N_1 is the number of spots on the visible disk at a given instant, regardless of spot size, N_2 is the number of groups of spots, and k is a coefficient which converts the results (averaged over a long time) from a given observatory to the results obtained
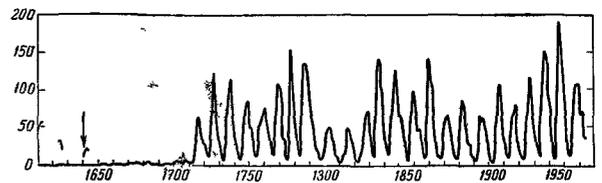
at the observatory in Zürich, adopted as a standard. Wolf plotted the time dependence of the daily values of W (now called *Wolf numbers*) over the years since 1818, the average monthly values since 1749, and the approximate average annual values since 1700. He also determined the years at which these values reached maxima and minima since the beginning of telescope measurements, in 1610, through 1699. The average length of the cycle in the oscillation of Wolf numbers (which we will call simply the *solar cycle*) turned out to be 11.2 yr, so that there have now been 20 complete cycles since the 1755 minimum. During each cycle, the Wolf number has varied from a few units to values of the order of 100–150 (here we might also note that the total area of the spots, A, averaged over a year, and expressed in millionths of the area of the visible hemisphere, i.e., in units of 3.02×10^{16} cm^2, is proportional to the Wolf number: $A=16.7W$). Figure 9 shows the Wolf numbers since 1610; we will be discussing this record further.

b) Spörer's law

The changes in the number of spots over the solar cycle are accompanied by changes in the distribution of spots over heliographic latitude. As early as 1858, Richard Carrington noted that the latitudes at which spots appear, in both the northern and sourthern hemi-spheres, decrease on the average during the solar cycle: The first spots of a cycle appear at the polar edges of the "royal latitudes," i.e., around 30 °N and 30°S, on the average; later spots usually appear progressively closer to the equator. At the maximum, they appear at 15°N and 15°S, and the last spots of a cycle appear near 8°N and 8°S. This behavior was soon confirmed by Wolf; it was rediscovered in 1867 by Gustav Spörer and P. A. Secchi; and in 1894 Spörer traced it back to 1621. The effect is sometimes called "Spörer's law." It can be demonstrated well by plotting all spots on a latitude vs time diagram (Fig. 10); this was first done in 1922 by Maunder, and the plots are referred to now as "Maunder butterfly diagrams." This diagram
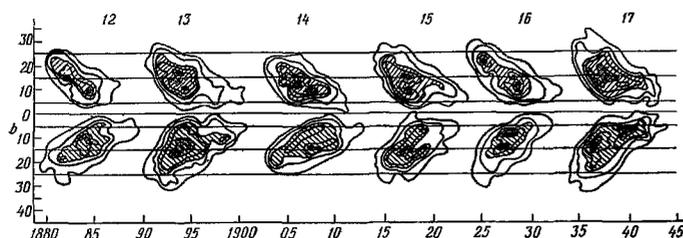

FIG. 10. Latitude distribution of sunspots over the period 1874–1913 ("Maunder butterflies").

shows, incidentally, that the first spots of a given cycle appear at the high latitudes before the last spots of the preceding cycle appear at low latitudes; this overlap of cycles, which amounts to 3 yr on the average, tends to smooth over the minima in the time dependence of the Wolf number. At the same time, adjacent cycles are independent in the sense that the corresponding "butterflies" do not intersect.

c) Hale–Nicholson laws

George Hale and S. Nicholson showed that within a solar cycle all the p spots in bipolar magnetic regions in one hemisphere and all the f spots in such regions in the other hemisphere have the same polarity, and in the next cycle the polarities of all such spots reverse, so that each solar cycle is an epoch of constant polarity of the solar magnetic field. A transition between cycles corresponds to a magnetic field reversal, and a complete magnetic cycle contains two successive spot cycles. These Hale–Nicholson laws show, on the one hand, that the mechanism which generates the solar magnetic field acts in an oscillatory fashion, causing quite regular (quasiperiodic) reversals of the field polarity; on the other hand, they show that the same mechanism is responsible for generating the solar magnetic field and the oscillation in the solar activity. It thus becomes clear that the theory of the solar cycle is a global problem of the magnetohydrodynamics of the sun.

The oscillation in solar activity can be seen not only in sunspots and magnetic fields but also in other phenomena in the solar atmosphere (Section 2), including the frequency of chromospheric flares and associated intensification of the x-ray, UV, and radio emission of the sun; the emission of corpuscular streams and cosmic rays; the number and distribution of protuberances (the major prominences zones shift toward the equator during the cycle, while the high-latitude prominence zones shift toward the poles, reaching them at the activity maxima); and the distribution of coronal streamers. [At activity minima, these rays are well-defined only at lower latitudes, and with distance from the sun they slope more and more toward the equator. In the polar regions, they form short, regular sheafs, which apparently correspond to magnetic lines of force. At activity maxima, on the other hand, when the polar prominences nearly reach the poles, the coronal streamers are identical in all directions, on the average.]

d) Waldmeier eruption hypothesis

Let us examine the irregularities observed as the solar cycle repeats itself. There are changes in the length of the cycle, in the shape of the curve giving the time dependence of the average annual Wolf number, $W(t)$, and in the maximum and minimum values of this number. Over the period 1755–1947, for example, the time intervals between successive minima varied from 9.0 to 13.6 yr; that between successive maxima varied from 7.3 to 17.1 yr (with a far greater scatter); the value of the minimum varied from 0.0 to 11.2; and the value of the maximum varied from 48.7 to 189.5. The

shape of the $W(t)$ curve is determined primarily by the value of W_{max}: The area under the rising branch of the W curve is nearly independent of W_{max} (so that large values of W_{max} mean short rise times and, in general, short cycles), while the area under the descending branch is proportional to W_{max}. While the spots are moving along the latitude, the scatter is smaller than during oscillations in the Wolf number. The latitudes at which the spots appear increase with increasing W_{max}.

The average annual value of W_{max} in a given cycle might be predictable from the value of W_{min} at the beginning of the cycle or, more precisely, from the extent to which the geomagnetic field is perturbed at the beginning of the cycle, as measured by the average annual value of the so-called AA index (the sum of the intervals over which the geomagnetic field changes in a 3-h time interval at two antipodal points on the earth's surface). Kane (1978) has shown that the average annual value of W_{max} correlates well with the average annual value of AA at the beginning of the cycle (the correlation coefficient is about 0.9).

All attempts to predict the characteristics of a solar cycle from those of the preceding cycle have failed. In 1935, M. Waldmeier offered an "eruption hypothesis," according to which solar cycles are completely independent of preceding cycles, as if each given cycle were a manifestation of some new eruption which has occurred within the sun. A clearer quantitative formulation of this property of the solar cycles was worked out by Gudzenko and Chertoprud (1964a and 1964b) [see also Gudzenko (1972)] through an interesting statistical analysis of data on the oscillations in the Wolf number $W(t)$.

e) Relaxation nature of the solar cycle

Gudzenko and Chertoprud plotted sliding annual average values of $W(t)$ over 19 cycles on the plane of $x = W, y = dW/dt$. They found an average closed curve $\mathscr{L} = \{ x = \bar{x}(t), y = \bar{y}(t) \}$ to describe the average solar cycle. For each point $M(t) = \{ x(t), y(t) \}$ they determined its projection $M'(t)$ on the curve $\mathscr{L}$ along the normal to this curve; the distance $n(t) = MM'$; and the phase $\theta(t) = t + \gamma(t)$ of the point $M'(t)$ on the $\mathscr{L}$ curve, so that $\gamma(t)$ is the distance along $\mathscr{L}$ from the point $M'(t)$ to the point $\bar{M}(t)$, which is moving uniformly along the average cycle $\mathscr{L}$. Now an attempt can be made to describe the trajectory $M(t)$ by means of the differential equations of an oscillator:

$$\frac{dn}{dt} = -N[\theta]n + F_n(t),$$
$$\frac{d\gamma}{dt} = \varkappa[\theta]n + F_\theta(t),$$
(4.1)

where $N[\theta]$ is the "stiffness" of the oscillator [it describes the stiffness of the "spring" which pulls the point $M(t)$ toward the cycle $\mathscr{L}$]; $\varkappa[\theta]$ is the "deviation from isochronism" for the oscillator [it describes the rate of change of the oscillation frequency as the point $M(t)$ moves away from the cycle $\mathscr{L}$]; and the functions $F_n(t)$ and $F_\theta(t)$ describe the fluctuation noise which perturbs the operation of the oscillator. For the simplest harmonic oscillator—a frictionless pendulum—the cycle $\mathscr{L}$ is a circle. During the oscillation, there is a peri-

odic conversion of potential energy into kinetic energy and back.' A periodic conversion of energy from certain forms into other forms is a general property of *conservative* self-excited oscillatory systems (here "conservative" means that the systems do not exchange energy with the external medium). For oscillators which are nearly conservative (for which the energy acquired from, or lost to, the external medium is small) most of the trajectories $M(t)$ are spirals which run repeatedly around the cycle $\mathscr{L}$, approaching it slowly (in comparison with the period of the cycle, τ). For these spirals, the average stiffness $\bar{N}$ is positive and small ($\bar{N} \ll 1/\tau$).

The opposite property is exhibited by "relaxation oscillators," in which energy is not converted from certain forms into others but is instead accumulated up to a certain level and then rapidly discharged from the system, so that upon an excursion from the cycle $\mathscr{L}$ the point $M(t)$ returns to it rapidly (in comparison with τ). Here the average stiffness is large ($\bar{N} \gg 1/\tau$). An example might be a water tank which water enters from the top at a constant velocity v and flows out the bottom through a small-aperture valve in drops of mass m. This mass is determined by surface tension alone and is independent of the water level U in the tank, while the drop-formation period does depend on U (and only in the steady state does the equality $\tau = m/v$ hold).

Gudzenko and Chertoprud studied the stiffness of the oscillator generating the solar cycle, finding, first, that it was independent of τ, confirming the validity of Eqs. (4.1) in this case, and second, that the average stiffness was large, $\bar{N} > 20/\tau$ {the largest values of $N[\theta]$ were found before the $W(t)$ minimum, where $N[\theta] > 60/\tau$}, so that the solar-cycle oscillator is clearly a relaxation oscillator: It quickly "forgets" early excursions from the average cycle $\mathscr{L}$, and this behavior is a quantitative formulation of the Waldmeier explosion hypothesis. It can thus be concluded that the variations in the solar activity are not a consequence of a periodic conversion of certain forms of energy into others (in particular, these variations are not a consequence of an energy transfer between the poloidal and toroidal components of the magnetic field or between kinetic energy and magnetic energy) and are instead a consequence of an accumulation of energy, followed by discharges of energy from a self-excited oscillatory system.

An estimate of the deviation from an isochronous behavior for the solar-cycle oscillator yielded a value of approximately zero, $\varkappa \approx 0$. We see from the second equation in (4.1) that the fluctuations in the phase of the solar cycle, $\gamma(t) = \int_0^t F_\theta(t) dt$, should behave by analogy with the coordinate of a free Brownian particle, so that the phase dispersion should increase over time in a "diffusion manner," $\langle \gamma^2(t) \rangle = 2Dt$. Data on the Wolf numbers over 19 cycles yield a statistically significant dependence of the phase dispersion on the time only for times $t < 4\tau$, no greater than four periods of the cycle. For such short times, Gudzenko and Chertoprud confirmed the diffusion law.

A far longer time interval can be spanned by making use of historical data on auroral observations (the num-
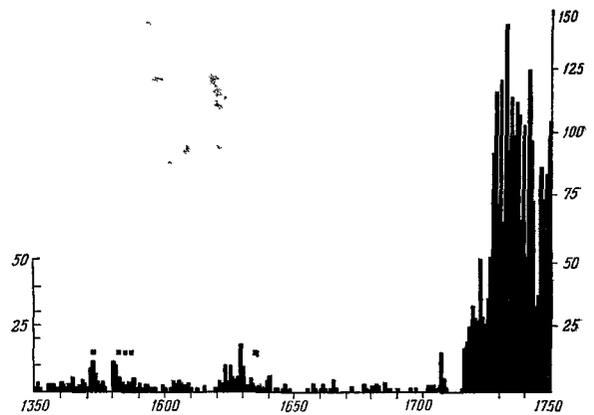

FIG. 11. Number of auroral observations over the period 1550–1750, according to Fritz. The small squares represent observations in the Far East.

ber of nights with auroral displays at intermediate latitudes correlates well with the Wolf number). The most complete catalog of such events was compiled in 1873 by H. Fritz (Fig. 11 is taken from his catalog for the period 1550–1750). In 1898, he published the dates of solar-activity maxima over the past 2000 yr corresponding to this catalog. In 1935, E. E. Slutskii noted an extremely long-term relationship between the fluctuations in the phase of the solar cycle (including a relationship between the lengths of cycles). Using Fritz' data to calculate the phase dispersion $\langle \gamma^2(t) \rangle$ over a long time interval, Gudzenko and Chertoprud found a curve described by

$$\langle \gamma^2(t) \rangle = \gamma_\infty^2 (1 - e^{-2t/\tau_0}), \qquad (4.2)$$

which corresponds to the dispersion in the coordinates of a bound Brownian particle (γ_∞ turned out to be larger than $\tau/4$). Accordingly, a phase feedback must also be introduced in Eq. (4.1). This is a general property of relaxation oscillators. In our example of drop formation, the phase feedback makes the drop-formation period dependent on the water level in the tank.

f) Maunder minimum

In addition to the irregularities from cycle to cycle in the solar activity, there are apparently some irregularities with far larger time scales and amplitudes. For example, exceedingly few sunspots were detected over the 70-yr period between 1645 and 1715 (Fig. 9). Spörer addressed this question in papers in 1887 and 1889, which were summarized in 1890 by Maunder. In 1894, in a paper on "A prolonged sunspot minimum," Maunder published the results of a further study of the reality of the effect. In 1922, he returned to this question in another paper with the same title. For a long time thereafter, however, this work received no further attention, until Eddy (1976 and 1977) published several more pieces of evidence showing the reality of the sharp decrease in solar activity in 1645–1715. Eddy called this period the "Maunder minimum" (it nearly coincided with the 1643–1715 reign of Louis XIV, the "Sun King").

According to the historical record, no sunspots at all were observed in 1656–1660, 1661–1671, 1689–1695,

1695–1700, and 1710–1713. During the Maunder minimum, the observation of sunspots was treated as a discovery warranting publication in a special paper. Discovering a spot in 1671, the well-known Parisian astronomer Giovanni Cassini wrote that for about 20 yr before this event astronomers had seen no significant sunspots. The same was reported by Jean Picard, John Flamsteed, Derham, and other astronomers, and these comments were published in several books on astronomy, including *Astronomy* by J. Lalande in 1792 and a book by Sir William Herschel in 1801. The situation was the same in observations of sunspots with the naked eye in the Far East: Of 143 such observations between 28 B.C. and 1743 A.D. which were published by Sigura Kanda in 1933, none fall in the Maunder minimum (according to this catalog, the longest period without visual observation of sunspots was 579–808, while the highest frequency of visual observations was in 1080–1280). Further confirmation of the reality of the Maunder minimum comes from catalogs of observations of auroras (see, for example, Fig. 11). For example, the first recorded auroral observation during the Maunder minimum was in 1708; the second, in 1716, served as the subject of a special paper by the Astronomer Royal Edmond Halley, who, 60 at the time, had never before observed an auroral display. We might also note that descriptions of the solar corona as it was observed during total eclipses of the sun contained nothing about coronal rays until 1715 (R. Cotes). All the earlier descriptions of the corona read like modern descriptions of the zodiacal light.

Yet another piece of evidence is the increase by more than 10% in the relative abundance of radiocarbon, C^{14}, in the annual rings of trees which occurred in 1640–1720, peaking about 1690, as discovered in 1958 by H. de Vries (and the effect is now called the "de Vries fluctuation"). This fluctuation has since been confirmed by extensive data from several countries. (C^{14} forms in atmospheric CO_2 as the result of bombardment by galactic cosmic rays, which reach a maximum intensity in years of a quiet sun, when their partial shielding by the magnetic fields of the solar wind is reduced. On the other hand, fluctuations in the C^{14} level in the atmosphere and thus in tree rings can result from other factors, for example, oscillations in the intensity of the geomagnetic field. Furthermore, in the twentieth century we have seen the "Suess effect": a decrease in the amount of C^{14} in atmospheric CO_2, due to a dilution by C^{14}-poor products of the combustion of fossil fuels.)

Data on the variations in the C^{14} relative abundance in annual tree rings have been summarized by Suess (1965) and Lerman *et al.* (1970), among others. Interpreting these data in solar-activity terms, Eddy (1976) distinguished in the variations over the past 1000 yr a "century-average maximum" in 1100–1250, a "Spörer minimum" in 1460–1550, a "Maunder minimum" in 1645–1715, and a modern maximum. Later (1977), Eddy went back even further, finding 12 extrema in the solar activity over the past 5000 yr on the basis of the C^{14} data: Sumerian, Pyramid, and Stonehenge maxima; Egyptian, Homeric, and Greek minima; a Roman maximum; a Medieval minimum; and the four extrema listed above.

Eddy *et al.* (1976) analyzed observations of the motion of spots over the solar disk which had been carried out by Christoph Scheiner in 1625–1626 and by Johannes Hevelius in 1642–1644, finding that for the 20 yr preceding the Maunder minimum the differential rotation of the sun was the same as at present; that the rotation at the equator accelerated by 3–4% at the beginning of this minimum; and that the difference between the rotation velocities at latitudes of 0° and 20° tripled. Howard (1976) has noted a similar acceleration (but only half as strong) in the rotation at the equator in years of quiet sun by examining the data from recent Doppler measurements. It may be that at activity maxima the equatorial rotation is retarded by Maxwell stresses in the magnetic field.

g) Babcock model

As mentioned earlier in connection with the Hale–Nicholson laws, a successful theory of the solar cycle would be expected to explain the mechanism which generates the global solar magnetic field and the quasiperiodic reversals of its polarity. The vortical part of the solar magnetic field is conveniently written in the form

$$\mathbf{H} = \mathrm{rot}\,(T\mathbf{r}) + \mathrm{rot}\,\mathrm{rot}\,(P\mathbf{r}), \qquad (4.3)$$

where the first term, which is perpendicular to the radius vector $\mathbf{r}$, is called the *toroidal* field (in the axially symmetric case, its vector lines are circles centered on the symmetry axis and lying in planes running perpendicular to this axis), and the second term is the *poloidal* field (in the axially symmetric case, its vector lines lie in meridional planes). In modern models of the heliomagnetic dynamo it is generally assumed that the toroidal field is created from the poloidal field through the differential rotation of the convection zone. Babcock was one of the first to advance a model of this type (1961).

According to this model, a poloidal magnetic field (with a total magnetic flux on the order of 8×10^{21} Mx) forms three years before the beginning of a solar cycle. The lines of force of this field run along meridians below the solar surface, at a certain depth in the convection zone, at low and intermediate latitudes (at, say, $|\varphi| < 55°$), while at polar latitudes ($|\varphi| > 55°$) the lines of force emerge from below the surface and close far above it, like the lines of force of a magnetic dipole. The differential rotation (with respect to latitude) ($\omega = 14°.28 - 2°.77 \sin^2\varphi$ per day, according to Newton and Nunn) stretches out the parts of the lines of force which lie inside the convection zone, causing a longitudinal displacement of these lines equal to $\delta\lambda = 17.6(n+3)\sin^2\varphi$ rad over the course of $n+3$ yr, and converting these lines into spirals, which make angles ψ with the meridian, where $\mathrm{tg}\,\psi = d\delta\lambda/d\varphi = 35.2(n+3)\times\sin\varphi\,\cos\varphi$. These spirals have an intensity $H = H_0 \sec\varphi \times \sec\psi \approx 35.2(n+3)H_0\sin\varphi$, where $H_0 \sim 5$ G is the intensity of the original poloidal field at the equator. In 3 yr, the extension ($n=0$) of the field at latitudes $|\varphi| = 30°$ reaches a critical value H_c (the average value of this

615   Sov. Phys. Usp. 23(9), Sept. 1980

A. S. Monin   615

critical field over the entire magnetized layer is H_c = 264 G, but it is several times higher in certain magnetic tubes). When this critical value is reached, spot formation begins at these latitudes (because the stretched parts of magnetic tubes rise), and further extension is stopped. At lower latitudes, the extension continues, so that in $n + 3$ yr the critical value H_c is reached at latitudes φ determined by the equation $|\sin\varphi| = 1.5/(n + 3)$, which describes Spörer's law in this model. The Hale–Nicholson polarity laws evidently hold also. Let us assume that the original "buried" poloidal field consists of, say, eight magnetic tubes with lengths of order $2R_\odot$; then in 3 yr they stretch out into spirals which girdle the sun five times in each hemisphere; they reach a total length of the order of $500R_\odot$, and eventually, over the entire solar cycle, they reach a length on the order of $10^3 R_\odot$. If each section of the extended tube of length $R_\odot$ is capable of generating three or four bipolar magnetic regions, then about 3000 such regions will form over the solar cycle, in good agreement with observational data.

Babcock's model goes on to the formation of a new poloidal field of opposite polarity, but here the arguments become far more hypothetical. Now it is argued that loops of magnetic force lines (with p segments closer to the equator than f segments, so that these loops now contain a contribution to the new poloidal field of opposite polarity) form above bipolar magnetic regions, rise high into the corona, and are pinched and carried off by the solar wind. The ends of the lines of force in the chromosphere reconnect, contributing to a new poloidal field, which then drops below the photosphere.

h) Leighton's equations

In a quantitative description of the MHD generation of a magnetic field, the evolution of the field is described by the induction equation

$$\frac{\partial H}{\partial t} - \text{rot}\,[\mathbf{V}H] = \nu_m \Delta H, \quad \text{div } H = 0, \quad (4.4)$$

where $\nu_m = c^2/4\pi\sigma$ is the magnetic viscosity coefficient (c is the speed of light, and σ is the elctrical conductivity of the medium). Taking the average of this equation, by analogy with (3.9)–(3.12), and assuming that the average magnetic field is a time-varying field, we find the following equations, in components along the axes of *a spherical coordinate system:*

$$\frac{\partial \bar{H}_r}{\partial t} - \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\,(V_r\bar{H}_\theta - V_\theta\bar{H}_r)$$
$$= \nu_m\left(\Delta\bar{H}_r - \frac{2}{r^2}\bar{H}_r - \frac{2}{r^2\sin\theta}\frac{\partial\sin\theta\bar{H}_\theta}{\partial\theta}\right)$$
$$+ \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\,(\overline{V'_r H'_\theta} - \overline{V'_\theta H'_r}), \quad (4.5)$$

$$\frac{\partial\bar{H}_\theta}{\partial t} + \frac{1}{r}\frac{\partial}{\partial r}r\,(\overline{V_r}\bar{H}_\theta - \overline{V_\theta}\bar{H}_r) = \nu_m\left(\Delta\bar{H}_\theta - \frac{\bar{H}_\theta}{r^2\sin^2\theta} + \frac{2}{r^2}\frac{\partial\bar{H}_r}{\partial\theta}\right)$$
$$- \frac{1}{r}\frac{\partial}{\partial r}r\,(\overline{V'_r H'_\theta} - \overline{V'_\theta H'_r}), \quad (4.6)$$

$$\frac{\partial\bar{H}_\lambda}{\partial t} - r\sin\theta\left(\bar{H}_r\frac{\partial\omega}{\partial r} + \frac{\bar{H}_\theta}{r}\frac{\partial\omega}{\partial\theta}\right) + \frac{1}{r}\left(\frac{\partial}{\partial r}r\bar{V}_r\bar{H}_\lambda - \frac{\partial}{\partial\theta}\bar{V}_\theta\bar{H}_\lambda\right)$$
$$\approx \nu_m\left(\Delta\bar{H}_\lambda - \frac{\bar{H}_\lambda}{r^2\sin\theta}\right) - \frac{1}{r}\left[\frac{\partial}{\partial r}r\,(\overline{V'_\lambda H'_r} - \overline{V'_r H'_\lambda}) - \frac{\partial}{\partial\theta}(\overline{V'_\theta H'_\lambda} - \overline{V'_\lambda H'_\theta})\right],$$
$$(4.7)$$

$$\frac{1}{r^2}\frac{\partial}{\partial r}r^2\bar{H}_r + \frac{1}{r\sin\theta}\frac{\partial}{\partial\theta}\sin\theta\bar{H}_\theta = 0. \quad (4.8)$$

Leighton (1969) suggested calculating the characteristics of the solar cycle from a semiempirical modification of Eqs. (4.5)–(4.8) adopted as a quantitative description of Babcock's model. Leighton's equations ignore the average meridional circulation, V_r, V_θ; the angular velocity of the differential rotation, ω, is given by

$$\omega = (\omega_0 + 18\sin^2\theta) + (\alpha + \beta\sin^n\theta)\frac{R_\odot - r}{h}R_\odot - h \leqslant r \leqslant R_\odot, \quad (4.9)$$

where the first expression in parentheses describes the differential rotation with respect to latitude according to Newton and Nunn (in radians per year), and the second term, with the adjustable parameters α, β, n, and h, describes the radial differential rotation. The equations are integrated over depth; in the case $\varepsilon = 0$, with h small in comparison with the thickness of the magnetized layer, it is assumed that $\bar{H} = 0$ and $\bar{H}_v = (\bar{H}_r)_{r=R_\odot}$ in the shear layer $R_\odot - h \leqslant r \leqslant R_\odot$. In the case $\varepsilon = 1$, on the other hand, with h equal to the thickness of the magnetized layer, $\bar{H}_\theta$ is assumed constant over depth, $\bar{H}_r \approx \frac{1}{2}(\bar{H}_r)_{r=R_\odot}$, and H_λ is assumed to vary approximately linearly with depth. After the averaging over depth, the equations become

$$\frac{\partial\bar{H}_r}{\partial t} = -\delta\frac{Fh}{80\tau R_\odot}\frac{\partial}{\partial\mu}\mu\bar{H}_\lambda + \frac{1}{T_D}\frac{\partial}{\partial\mu}(1 - \mu^2)\frac{\partial\bar{H}_r}{\partial\mu}, \quad (4.5')$$

$$\frac{\partial\bar{H}_{r0}}{\partial t} = -\frac{Gh}{80\tau R_\odot}\frac{\partial}{\partial\mu}\mu\bar{H}_\lambda - \frac{\bar{H}_{r\theta}}{50}; \quad (4.5'')$$

$$\frac{\partial\bar{H}_\lambda}{\partial t} = \sin\theta\left[-(\alpha + \beta\sin^n\theta)\frac{R_\odot}{h}(\bar{H}_r + \bar{H}_{r0})\right.$$
$$\left. + \varepsilon\left(36 + \frac{n\beta}{2}\sin^{n-2}\theta\right)\sin\theta\cos\theta\,\bar{H}_\theta\right] - \delta\frac{|\bar{H}_\lambda|\bar{H}_\lambda}{100H_c\tau} - \frac{\bar{H}_\lambda}{50}, \quad (4.7')$$

$$H_\theta = \frac{R_\odot}{h\sin\theta}\int_{-1}^{\mu}(\bar{H}_r + \bar{H}_{r0})\,d\mu. \quad (4.8')$$

Here the time t is expressed in years; $\mu = \cos\theta$; and the factor δ, which is zero at $|\bar{H}_\lambda| \leqslant H_c$ and one at $|H_\lambda| > H_c$, describes the ascent of the sections of the magnetic tubes of the toroidal field which have been stretched above the critical intensity H_c. The first terms on the right sides of (4.5')–(4.5'') describe the increase in the radial field during the ascent of the stretched parts of the toroidal field, and they describe the formation of groups of spots of length a (the empirical value is $a/4\pi R_\odot \approx 1/80$) and making angles γ with the parallels which satisfy ($\sin\gamma \approx \mu/2$). Here F is a correction factor, which may differ from unity because of an inaccurate specification of a and γ (clearly, as oscillatory solution is possible only if $F \geqslant F_{min}$, and F_{min} must not be much greater than unity). The time τ specifies the relative rate of formation of the ascending regions. The second term on the right side of (4.5) describes the meridional diffusion of the radial field which is caused by the supergranules (the scale time for this diffusion is $TD = 20$ yr). If we wish to avoid the possibility of a complete damping of weak fields (weaker than H_c) by this diffusion, we single out a small part H_{r0} of the radial field (representing a fraction $G \approx 0.003F$), which does not diffuse (but which decays with a scale time of 50 yr). The first term on the right side of Eq. (4.7) describes the formation of the toroidal field from the poloidal field as a result of the differential rotation; the second describes the weakening of the toroidal field due to the ascent of some of its stretched regions (this

weakening is $a\bar{H}_\lambda/2\pi R_\odot \tau$; with $a=a_0|H_\lambda|/H_c$ and $a_0/2\pi R_\odot \approx 1/100$). The third term describes field decay with a scale time of 50 yr.

Equations (4.5')–(4.8') were integrated for a given field $\bar{H}_{r0}(\mu)$ and for $\bar{H}_r=\bar{H}_\lambda\overset{\cdot}{=}0$ at $t=0$. The value adopted for F was the minimum value at which an oscillatory situation could still be attained, and τ was chosen in each step in such a manner that the oscillation period approached 22 yr. It turned out that, even in the absence of a radial differential rotation ($\alpha=\beta=0$; here $\varepsilon=1$, $F=10$, $\tau=0.42$, and $H_c=20R_\odot/h$), the curves of constant level of the field $\bar{H}_\lambda(\theta,t)$ (and also those of the fields $\bar{H}_r$ and $\bar{H}_\theta$) are extremely similar to Maunder butterflies. Here, however, the value $F_{min} \approx 6$ is too large. To reduce F_{min} we must choose positive values for α or β (i.e., we must assume that ω *increases with depth*). With $\alpha=0$ and $\beta=18$ (the change in ω across the shear layer at the equator is the same as that between the equator and the pole at the surface), the value of F_{min} turns out to be of the order of 0.6 (so that, in terms of the value of F_{min}, the radial gradient in ω is ten times more effective than the latitudinal gradient). With $\beta=-18$ and $n=2$, the bottom of the shear layer is in a state of rigid rotation, and with $\alpha=18$ the period is the same as that for the equator at the surface. At very large values of α, the period is much smaller, but in such cases the Maunder butterflies lie at latitudes which are too high. The best quantitative agreement with Spörer's law comes from the "standard" model with $\alpha=0$, $\beta=10$, $n=8$, $\varepsilon=1$.

In one of his numerical simulations, Leighton randomized the standard model, choosing the value of τ at each latitude from a log-normal distribution three times each year; this distribution provided a mean period of 11 yr for the solar cycle and a standard distribution of 1.6 yr in this period ($F=2$, $\tau_0=0.6$, and $\sigma_{int}=1$). This method was used to calculate 20 successive solar cycles. Their statistical characteristics turned out to agree well with the actual data.

i) Hydromagnetic dynamos

The most important term in Leighton's equations is the first term on the right side of (4.5'), which describes the generation of a poloidal magnetic field by a toroidal emf:

$$F_\lambda = |V'H'|_r = \bar{a}H_r, \quad a \sim \frac{\delta Fh}{40\tau R_\odot}\frac{\sin\gamma}{\sin\theta}. \tag{4.10}$$

Equations for a hydromagnetic dynamo with a toroidal emf of the type $F_\lambda = \alpha\bar{H}_\lambda$ (the "α effect") were first derived for the model of fine-scale "cyclonic convection" by Parker (1955) [see also Parker (1970, 1971)]; they were subsequently derived for slow, large-scale convection by Braginskii (1964a, 1964b, 1965). Analogous equations can be derived for a turbulence which does not have reflection symmetry and which thus has a nonzero *spirality*, $\overline{V\,rot V}\neq 0$. In the field of this turbulence, an emf $E=\alpha H - \beta\,rotH$ forms, where $\alpha=-1/3\,\tau\overline{V\,rotV}$ and $\beta=(1/3)\tau\overline{VV}$. Here τ is the correlation scale time of the turbulence. The effectiveness of this dynamo mechanism was demonstrated by M. Steenbeck, F. Krause, and K. Rädler at Potsdam [see, for exam-

ple, the review by Krause and Rädler (1971)] and also by Moffatt (1970). Work on the heliomagnetic dynamo is reviewed by Vandakurov (1976) and Stix (1976). In many of the papers, the authors restrict the problem to the "kinematic formulation," in which the velocity field **V** is given, rather than calculated, and the magnetic field is calculated from the induction equation with the α effect, e.g.,

$$\frac{\partial H}{\partial t} = rot\,|VH| - \alpha H - \gamma\,rot\frac{H}{\mu}, \tag{4.11}$$

where μ is the magnetic permeability of the plasma, which is equal to unity outside the convection zone and satisfies $\mu\ll 1$ in this zone [this diamagnetic property of a turbulent plasma was discovered by Zel'dovich (1956), who showed for the case of a two-dimensional turbulence that a large-scale magnetic field is "ejected" from turbulence regions into regions with less intense turbulence. The ejection of the magnetic lines of force should increase the scale times for changes in the magnetic fields in the convection zone, eliminating the difficulty of finding too short a period for the solar cycle if the diamagnetic effect is ignored.

As examples, we can cite the numerical simulations carried out using (4.11) by Ivanova and Ruzmaĭkin (1976, 1977), in which ω, α, γ, and μ are assigned simple functional dependences on the depth [in the nonlinear model of 1977 it was assumed that $\alpha=\alpha_0(r)(1+\xi H^2)^{-1}$, where the coefficient ξ depends on the intensity of the spiral part of the turbulent velocity field and determines the amplitude of the steady-state magnetic field]. These simulations showed that the variations in the magnetic field over the cycle are wavelike [Parker (1955) described them as a "dynamo wave"]. The direction in which this wave propagates depends on the sign of the product $\alpha\partial\omega/\partial r$: Spörer's law and the correct shape for the Maunder butterflies are found if $\alpha\partial\omega/\partial r<0$ in the northern hemisphere [analysis of the phase relations for the oscillations in the poloidal and azimuthal fields by the method of Stix (1976) yields $\partial\omega/\partial r<0$]. A steady-state cycle is found only for a certain value, D_0, of the dimensionless dynamo number $D=(\alpha_0/\beta_0^2)\partial\omega_0/\partial r$. At $D<D_0$, the oscillation is damped; at $D>D_0$, it grows, and in order to stop the growth it is necessary to introduce an inverse effect of the magnetic field on the motion, i.e., a nonlinearity. The simplest way to introduce a nonlinearity is to allow a decrease in α with increasing H (the magnetic field should first and foremost suppress the factor primarily responsible for its own increase: the spirality), as recommended by Vaĭnshtein and Zel'dovich (1972). The introduction of a nonlinearity opens up new possibilities, for example, finding an explanation for the long-period variations in the solar cycle, such as the Maunder minimum [see Yoshimura (1978)].

Gilman (1968, 1969a, 1969b) has discussed a model of a heliomagnetic dynamo in which the elements of the spiral turbulence are Rossby waves which are excited in the convection zone by a presumed latitudinal temperature gradient. The vertical motions in the Rossby waves produce large-scale vertical fields from the toroidal magnetic field, and the Rossby waves transport these vertical fields to the poles, creating a poloidal field.

The differential rotation (again produced by Rossby waves) uses this poloidal field to form a new toroidal field, of the opposite sign.

However, Gilman carried out corresponding numerical simulations only for an extremely simplified model for the convection zone: a cylindrical annular gap with solid, electrically conducting walls, instead of a spherical shell. Gilman used a two-layer model instead of a continuous stratification, and he used very few Fourier components for the MHD fields. The calculations yielded quasiperiodic polarity reversals in the solar magnetic field, but the quantitative results were quite different from observations: The period was about 2 yr instead of 22 yr, and the maximum poloidal magnetic field in the polar regions was about 40 G, instead of 1–2 G. These quantitative discrepancies may result from the extremely simplified model or from the fact that something different is happening in the sun: Instead of ordinary Rossby waves, excited through the β effect in the presence of a latitudinal temperature gradient, there may be giant convection cells, twisted by the Coriolis force. These cells could form due to a vertical temperature gradient in an unstably stratified convection zone.

In summary, a theory for the solar cycle can probably be constructed through numerical simulations on the MHD of a set of giant convection cells. It will not be a trivial task to carry out these simulations without the simplifications adopted by Gilman, i.e., by using the actual geometry and a model with a large number of layers and a large number of Fourier components for the MHD fields. On the other hand, this problem is completely comparable to the problem of simulating the overall circulation in the earth's atmosphere, which has been pursued successfully at several research centers. The outlook for the solar problem is accordingly quite favorable.

L. H. Aller, The Abundance of the Elements, Interscience, New York, 1961 (Russ. Transl. IL, M., 1963).
L. H. Aller and D. B. McLaughlin (editors), Stellar Structure, U. Chicago Press, Chicago, 1965 (Russ. Transl. Mir, M., 1970).
H. W. Babcock, Astrophys. J. 133, 572 (1961).
L. Biermann, Z. Astrophys. 28, 304 (1951).
L. Biermann, in: Electromagnetic Phenomena in Cosmical Physics: IAU Symposium, No. 6, Stockholm, 1958, p. 248.
K. H. Böhm, in: IAU Symposium, No. 28, 1967, p. 366.
S. I. Braginskiĭ, Zh. Eksp. Teor. Fiz. 47, 2178 (1964a) [Sov. Phys. JETP 20, 1462 (1965)].
S. I. Braginskiĭ, Geomagn. Aeronom. 4, 732 (1964b).
J. C. Brandt and J. Heise, Astrophys. J. 159, 1057 (1970).
J. B. Brooker, G. R. Isaak, and H. B. Van der Raay, Nature 259, No. 5539, p. 92 (1976).
V. Bumba and R. Howard, Astrophys. J. 141, 1492 (1965a).
V. Bumba and R. Howard, Astrophys. J. 141, 1502 (1965b).
W. F. Cocke, Astrophys. J. 150, 1041 (1967).
R. P. Davies–Jones and P. A. Gilman, Solar Phys. 12, 3 (1970).
R. P. Davies–Jones and P. A. Gilman, J. Fluid Mech. 46, Part I, 65 (1971).
C. de Jager, "Structure and dynamics of the solar atmosphere," Handbuch der Physik, Vol. LII, Astrophysics III: The Solar System (ed. S. Flügge), Springer–Verlag, Berlin, 1959, p. 80 (Russ. Transl. IL, M., 1962).
R. H. Dicke and H. M. Goldenberg, Phys. Rev. Lett. 18, 313 (1967).
B. R. Durney, in: Basic Mechanisms of Solar Activity, IAU, 1976, p. 243.
B. R. Durney and I. W. Roxburgh, Solar Phys. 16, 3 (1971).
J. A. Eddy, Science 192, 1189 (1976).
J. A. Eddy, Sci. Am. 236, No. 5, 80 (1977).
J. A. Eddy, P. A. Gilman, and D. E. Trotter, Solar Phys. 46, 3 (1976).
K. Elsässer, Z. Astrophys. 63, 65 (1965).
P. A. Gilman, Science 160, 760 (1968).
P. A. Gilman, Solar Phys. 8, 316 (1969a) (I).
P. A. Gilman, Solar Phys. 9, Part 1, 3 (1969b) (II).
P. A. Gilman, Solar Phys. 27, 3 (1971).
P. A. Gilman, Part 2: Nonlinear Results, J. Fluid Mech. 57, Part 2, 381 (1973).
P. A. Gilman, Ann. Rev. Astron. Astrophys. 12, 47 (1974).
P. A. Gilman, J. Atm. Sci. 32, 1331 (1975).
P. A. Gilman, in: Basic Mechanisms of Solar Activity: Proc. of IAU Symposium, 1976, p. 207.
P. A. Gilman, Geophys. Astrophys. Fluid Dynamics 8, 93 (1977) (I).
P. A. Gilman, "Effects of temperature boundary conditions," Geophys. Astrophys. Fluid Dynamics 11, 157 (1978a) (II).
P. A. Gilman, "Effects of velocity boundary conditions," Geophys. Astrophys. Fluid Dynamics 11, 181 (1978b) (III).
G. S. Golitsyn, Astron. Zh. 49, 360 (1972) [Sov. Astron. 16, 294 (1972)].
G. S. Golitsyn, Vvedenie v dinamiku planetarnykh atmosfer (Introduction to the Dynamics of Planetary Atmospheres), Gidrometeoizdat, Leningrad, 1973.
L. I. Gudzenko, V poiskakh prirody solnechnykh pyaten (Seeking the Nature of Sunspots), Znanie, Moscow, 1972.
L. I. Gudzenko and V. V. Chertoprud, Astron. Zh. 39, 758 (1962) [Sov. Astron. 6, 590 (1962)].
L. I. Gudzenko and V. E. Chertoprud, Astron. Zh. 41, 697 (1964) [Sov. Astron. 8, 555 (1965)].
R. T. Hansen, S. F. Hansen, and H. G. Loomis, Solar Phys. 10, 135 (1969).
H. A. Hill, P. D. Clayton, D. L. Patz, A. W. Healy, R. I. Stebbins, J. R. Oleson, and C. A. Zanoni, Phys. Rev. Lett. 33, 1497 (1974).
R. Howard, Ann. Rev. Astron. Astrophys. 5, 1 (1967).
R. Howard, Solar Phys. 16, 21 (1971).
R. Howard, Sci. Am. 232, 106 (1975).
R. Howard, Astrophys. J. 210, Part 2, L159 (1976).
R. Howard and J. Harvey, Solar Phys. 12, 23 (1970).
T. S. Ivanova and A. A. Ruzmaĭkin, Astron. Zh. 53, 398 (1976) [Sov. Astron. 20, 227 (1976)].
T. S. Ivanova and A. A. Ruzmaĭkin, Astron. Zh. 54, 846 (1977) [Sov. Astron. 21, 479 (1977)].
R. P. Kane, Nature 274, 139 (1978).
S. P. Khromov, Meteorol. Gidrol. No. 9, 93 (1973).
R. Kippenhahn, Astrophys. J. 137, 664 (1963).
H. Köhler, Solar Phys. 13, 3 (1970).
F. Krause and K. H. Rädler, in: Solar Magnetic Fields: IAU Symposium, Dordrecht, Holland, D. Reidel, 1971, p. 770.
G. P. Kuiper (editor), The Solar System. Vol. 1., The Sun, U. Chicago Press, Chicago, 1953 (Russ. Transl. IL, M., 1957).
A. I. Lebedinskiĭ, Astron. Zh. 18, 10 (1941).
R. B. Leighton, Astrophys. J. 156, Part 1, 1 (1969).
J. G. Lerman, W. G. Nook, and J. C. Vogel, in: Radiocarbon Variations and Absolute Chronology, Almquist and

Wiksell, Stockholm, 1970, p. 275.

R. Lüst, in: Kosmicheskaya gazodinamika (Cosmic Gas Dynamics, Wiley) (Russ. Transl. Mir, M., 1972, p. 288).

H. K. Moffatt, J. Fluid Mech. 41, 435 (1970).

A. S. Monin, Izv. Akad. Nauk SSSR. Ser. Fiz. Atmos. Okeana 1, 45 (1965a).

A. S. Monin, Izv. Akad. Nauk SSSR. Ser. Fiz. Atmos. Okeana 1, 258 (1965b).

A. S. Monin, Izv. Akad. Nauk SSSR. Ser. Fiz. Atmos. Okeana 1, 490 (1965c).

A. S. Monin, Prognoz pogody kak zadacha fiziki (Weather Prediction as a Problem in Physics), Nauka, Moscow, 1969.

É. R. Mustel', Zvezdnye atmosfery (Stellar Atmospheres), Fizmatgiz, Moscow, 1960.

E. N. Parker, Astrophys. J. 122, 293 (1955).

E. N. Parker, Astrophys. J. 128, 664 (1958).

E. N. Parker, Astrophys. J. 160, No. 2, Part 1, 383 (1970).

E. N. Parker, in: The Solar and Terrestrial Dynamos, Astrophys. J. 164, No. 3, Part 1, 491 (1971).

S. B. Pikel'ner, Usp. Fiz. Nauk 88, 505 (1966) [Sov. Phys. Usp. 9, 236 (1966)].

A. B. Pittock, Rev. Geophys. Space Phys. 16, 400 (1978).

I. W. Roxburgh, in: Stellar Rotation. Proc. IAU Colloquium, Ohio State University, Columbus, September 8–11, 1969, p. 318.

A. A. Ruzmaĭkin, Priroda No. 3, 76 (1977).

A. A. Ruzmaĭkin and S. I. Vaĭnshteĭn, Astrophys. Space Sci. 59, 215 (1978).

K. H. Schatten, Solar Phys. 32, 315 (1973).

R. L. Sears, Astrophys. J. 140, 477 (1964).

A. V. Severny, V. A. Kotov, and T. T. Tsap, Nature 259, No. 5539, 87 (1976).

V. P. Starr, Physics of Negative Viscosity Phenomena, McGraw–Hill, New York, 1968 (Russ. Transl. Mir, M., 1971).

V. P. Starr and P. A. Gilman, Tellus 17, 334 (1965a).

V. P. Starr and P. A. Gilman, Astrophys. J. 141, 1119 (1965b).

V. P. Starr and P. A. Gilman, Sci. Am. 218, No. 1, 100 (1968).

M. Stix, Astron. Astrophys. 47, 243 (1976).

H. E. Suess, J. Geophys. Res. 70, 5937 (1965).

S. I. Vaĭnshteĭn and Ya. B. Zel'dovich, Usp. Fiz. Nauk 106, 431 (1972) [Sov. Phys. Usp. 15, 159 (1972)].

Yu. V. Vandakurov, Konvektsiya na Solntse i 11-letniĭ tsikl (Convection in the Sun and the 11-Year Cycle), Nauka, Leningrad, 1976.

S. V. Vorontsov and V. N. Zharkov, Astron. Zh. 55, 84 (1978) [Sov. Astron. 22, 46 (1978)].

F. Ward, Pure Appl. Geophys. 58, 157 (1964).

F. Ward, Pure Appl. Geophys. 60, 126 (1965a).

F. Ward, Astrophys. J. 141, 534 (1965b).

F. Ward, Astrophys. J. 145, 416 (1966a).

F. Ward, Pure Appl. Geophys. 63, 196 (1966b).

F. Ward, Mon. Not. R. Astron. Soc. 135, 147 (1976).

J. Wasiutynski, Astrophysica Norvegica (Oslo) 4, 497 (1946).

J. M. Wilcox and R. Howard, Solar Phys. 13, 251 (1970).

H. Yoshimura, Astrophys. J. 226, 706 (1978).

Ya. B. Zel'dovich, Zh. Eksp. Teor. Fiz. 31, 154 (1956) [Sov. Phys. JETP 4, 460 (1957)].

J. B. Zirker, Rev. Geophys. Space Phys. 15, 257 (1977).

Translated by Dave Parsons