AMERICAN JOURNAL of PHYSICS

Hydromagnetism. II. A Review

Walter M. Elsasser

Citation: Am. J. Phys. **24**, 85 (1956); doi: 10.1119/1.1934152
View online: http://dx.doi.org/10.1119/1.1934152
View Table of Contents: http://ajp.aapt.org/resource/1/AJPIAS/v24/i2
Published by the American Association of Physics Teachers

Hydromagnetism. II

A Review

Walter M. Elsasser*
University of Utah, Salt Lake City, Utah

(Received August 15, 1955)

The theoretical foundations having been given in Part I, phenomena here reviewed are: character and magnitude of the geomagnetic secular variation, the geomagnetic westward drift, hydromagnetic interpretation of irregularities in the earth's angular velocity; nature of fossil magnetism, magnetization of historical and prehistorical bricks and clays, of eruptive and sedimentary rocks, reversals of the earth's field in the geological past, evidence for pole migration; exploration of solar magnetic fields, the solar convection zone, magnetic fields of sunspots, the sunspot cycle, the over-all solar field, observations of magnetic stars, variable magnetic stars; fields in the solar corona, streamers and prominences, magnetic storms and auroras, planetary fields, the galactic field, the problem of cosmic rays.

GEOMAGNETIC SECULAR VARIATION

In Part I we have given an outline of the dynamo theory of the earth's magnetic field. This theory needs few assumptions; essentially it suffices to assume the existence of convection in the fluid, metallic core of the earth. The characteristics of the core are well known from seismic and other geophysical data[1] (the radius of the core is 0.55 of the earth's radius). The convective motions have usually been assumed to be driven by radioactivity, but Urey[2,3] has demonstrated that radioactive concentration in the core is most unlikely; one would thus prefer to use the energy liberated by such processes as slow chemical separations and crystallizations. In any event, the power required to drive the convection which in turn drives the dynamo mechanism is relatively small and presents no serious geophysical problems. Even if there was no loss of heat through the mantle[4] at all, and if the limitations of the available energy by the second law are taken into account, a source of energy that would continously maintain convection and the magnetic field need not heat up the core by more than a few hundred degrees throughout the lifetime of the earth. These are rather modest requirements. The main problems of the earth's magnetic field are distinctly dynamical. We are dealing with a system in which there prevails thermally induced convection in the presence of very strong Coriolis forces and ponderomotive forces of the magnetic field. Considering that after nearly a century of dynamical meteorology our understanding of a so readily observable phenomenon as the general circulation of the atmosphere is still quite rudimentary, it is little wonder that the equally complicated problems of the dynamics of the earth's core are not amenable to facile solutions. Nevertheless, the outline of the dynamo theory seems well established as was shown in Part I. The writer has reviewed the general geophysical setting of geomagnetism at considerable length elsewhere,[1] and we shall not renew here the discussion of these specialized questions.

The sum total of geomagnetic phenomena at the earth's surface may be divided into three parts: first, the main dipole, only immediate witness of the field-generating dynamo mechanism inside the core; secondly, the higher spherical-harmonic components of the field of internal origin beginning with the quadrupole terms (they do not seem to have a stationary part but are in their totality subject to secular variation); thirdly, the field of external origin, ionospheric variations as well as the important phenomenon of magnetic storms (these will be discussed in the last section).

Observations of the geomagnetic secular variation, sporadic, it is true, go a long way back into history. The role of Gauss in systematizing them

---

* Supported by the Office of Naval Research.

[1] W. M. Elsasser, Revs. Modern Phys. 22, 1 (1950).

[2] H. C. Urey, *The Planets* (Yale University Press, New Haven, 1952).

[3] H. C. Urey, Ann. géophys. 11, 65 (1955).

[4] *Mantle* is a generally accepted geophysical term for the outer, solid part of the earth surrounding the core.

in terms of developments in spherical harmonics is well known, but not until the end of the nineteenth century was a methodical coverage of most of the earth's surface achieved. As a result mainly of the observational activity of the Carnegie Institution, first under the direction of L. A. Bauer and later under J. A. Fleming,[5,6] we possess rather extensive world-wide records of the secular variation extending over several decades. This material has been exhaustively evaluated, mainly in the form of maps of the secular variation, by Vestine.[7] The subjective impact of these maps as first-hand evidence of the intensity and rapidity of the geomagnetic secular variation outweighs many mathematical demonstrations. We are not reproducing them here because they take up much space and we have reproduced some samples elsewhere[1] at a readily accessible place. The available evidence indicates[8] that the sources of the geomagnetic secular variation are below the earth's mantle. This conclusion is, moreover, plausible on general grounds since changes of that rapidity could hardly take place in a solid. Furthermore, the secular variation which we observe originates most likely in the top strata of the core, in a layer whose depth may be estimated as of the general order of 50–100 km. The reason for our inability to observe variations occurring at greater depth in the core is found in the screening effect of the superimposed layers of electrically conducting fluid. The mathematical treatment of this phenomenon follows the familiar theory of the skin effect. As compared with laboratory examples of skin effect, the shielding thicknesses are larger in proportion as the periods are longer.

The secular variation may be thought of as resulting from the perturbation of the main field (poloidal plus toroidal) by the local large-scale "eddies" of the fluid motion near the surface of the core. Vestine's maps[1,7] show that the secular variation has the appearance of a series of positive and negative "hills" spread irregularly over the earth. It should be noted, by the way, that these hills represent $\partial B/\partial t$, not B itself; the irregularities of B are more difficult to establish than those of its time derivative and can be given only in relatively coarser outline. Detailed studies of certain individual regions in regard to their geometry and kinematics have been made,[9,10] and a projection of the entire secular-variation field from the surface of the earth to the surface of the core (assuming zero conductivity in the mantle) was recently carried out by McDonald.[11] To understand what is implied in these investigations, consider the inverse problem of using a potential function defined on the surface of a sphere (with all sources inside) to determine the corresponding potential on a sphere of twice the size. Clearly, there will be a general smoothing out of local features owing to the outward projection. Quantitatively, it may be shown[11] that details less than 20–25 degrees apart are obliterated on the outward projection. The situation of an observer at the earth's surface is somewhat comparable to that of a man who has a photograph taken with a wavelength of the same order as the structural details of the object photographed. He might represent his object by so many individual scatterers whose locations and relative intensities he can roughly determine whereas their detailed structure eludes him. The hills and dales of the secular variation at the surface indicate the existence of much steeper features at the core whose precise details cannot be known. It has been possible[11] to represent one of Vestine's maps accurately by means of 65 hills of standardized shape and different strengths.

The analysis of the irregular and secular-variation field in terms of spherical surface harmonics is of some interest. The actual analysis has been carried through[1,7] up to and including terms with $n=6$. If a great-circle distance of 25° is the limit of resolution for projection down to the core, then the field at the earth's surface must be smooth enough so that in its spherical-harmonic series the coefficients go to zero in the neighborhood of $n=15$. At present the data are a long way from an accuracy where one could push the analysis that far. Owing to the preponderance of higher spherical harmonics the irregular and

---

[5] J. A. Fleming, Proc. Phys. Soc. (London) **58**, 213 (1946).

[6] J. A. Fleming, editor, *Terrestrial Magnetism and Electricity* (Dover Publications, New York, 1949).

[7] E. H. Vestine, Carnegie Inst. Wash., Publ. **578**, **580** (1947, 1948).

[8] W. M. Elsasser, Phys. Rev. **60**, 876 (1941).

[9] E. C. Bullard, Monthly Notices Roy. Astron. Soc., Geophys. Suppl. **5**, 248 (1948).

[10] F. Lowes and S. Runcorn, Trans. Roy Soc. (London) **243**, 525 (1951).

[11] K. McDonald, J. Geophys. Research (to be published).

secular-variation field increases much more rapidly as we go downwards into the earth than the dipole field. Quantitatively, the rms value of the irregular field (quadrupoles and above) at the earth's surface may be estimated from the data as about 5% of the rms dipole field.[1] If we assume that the center of gravity of the spherical-harmonic series for the irregular field is near the terms with $n=6$ (which seems a sufficiently conservative estimate) then on going to the core the irregular field has increased by a factor $2^4=16$ above the increase of the dipole field. At the surface of the core the rms irregular field will then be of the same magnitude as the mean dipole field. One must not conclude that the main dipole is merely the lowest harmonic component of a generally irregular field at the core. The direction of the main dipole axis (inclined by about $11\frac{1}{2}°$ relative to the geographical axis) is relatively stable and does in particular not partake of the general westward drift (see below) characteristic of all the harmonics above $n=1$. Physically speaking, this indicates in accordance with the general notions of the dynamo theory that the dipole originates farther down in the core, as distinct from the boundary-layer phenomena expressed by the irregular local fields and their variation.

So far we have spoken mainly about the space dependence of the geomagnetic irregularities. The secular-variation field expresses directly the time dependence of the irregular part of the geomagnetic field. Figure 1 shows a typical record of a magnetic observatory (the points representing yearly averages of the three force components and of the total force). The unit is the gamma ($1\gamma=10^{-5}$ gauss). Often the field changes very rapidly over a few years. This is apparently caused by localized "hills" moving about at the surface of the core and passing on occasion close by the observing station. Moreover, as Fig. 1 shows, the second derivative of the field with respect to time can be quite pronounced, corresponding to a large change of the first derivative in the course of several years. This sets an upper limit to the conductivity of the mantle through which these field variations have to pass on their way up from the core.

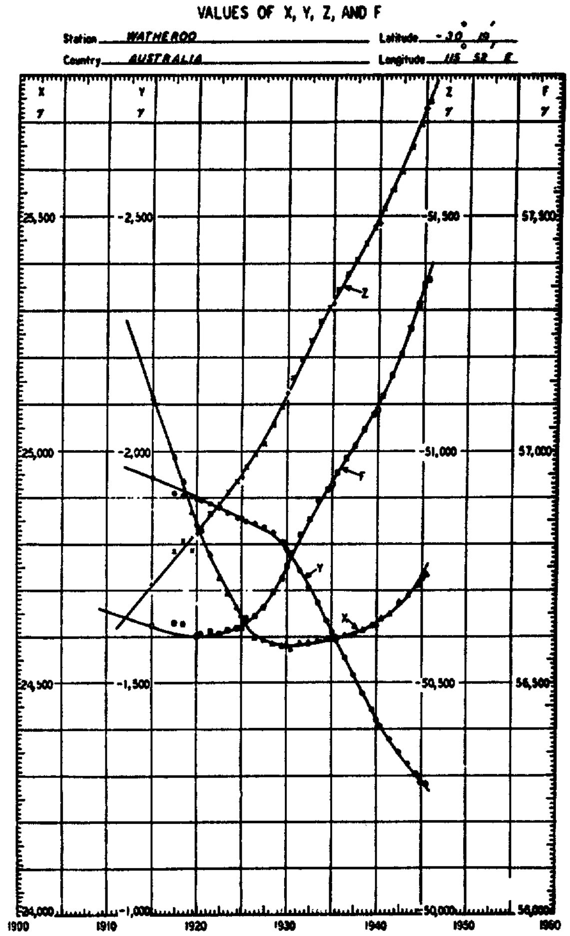


Fig. 1. Observatory record showing large second time derivatives (from Vestine's catalog).

Vestine's maps[7] show that the mean secular variation is of the order of 50γ/year with occasional high values of 150γ/year and slightly beyond. If we let such a mean variation act over 40 years we obtain a value of $2000\gamma=0.02$ gauss, a number fairly representative of the mean irregular field. Again, periods of 40 years are fairly representative of the secular variation. The view that the irregular field, from the quadrupoles on up, is in its totality subject to variation and has no stationary component, is thus consistent with the observations. According to the theory these irregular field components are created in the top layers of the core. It is a tempting problem to try to relate these variable fields quantitatively to fluid eddies near the boundary of the core. Calculations carried out by associates of the writer[12] have shown that the mathematics involved is

---

[12] W. L. Bade, unpublished report, University of Utah, 1954.

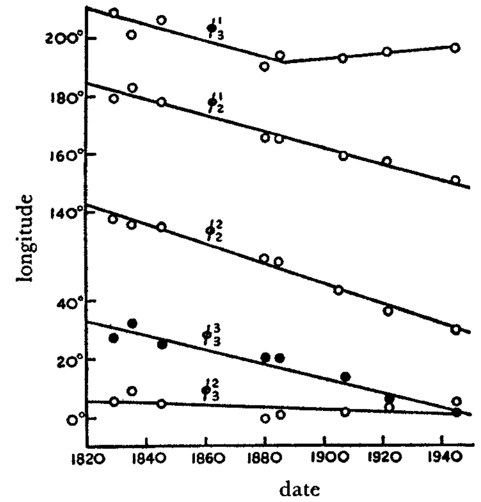


FIG. 2. Showing east-westerly drift of tesseral harmonics. (After Bullard and co-workers.)

very difficult, especially owing to the nature of the boundary conditions. It has not yet been possible to establish a close correspondence between eddies of the fluid motion and peaks of the secular variation.

The most intriguing feature of the geomagnetic secular variation is the so-called *westward drift*. The hills and dales of the secular variation not only grow and decline and shift about randomly, but also, superposed upon this irregular behavior, there is an average motion of the whole pattern from east to west at the rate of 0.18°/year. This figure was determined by Bullard and associates[13] who have analyzed the phenomenon in much detail. Figure 2, from their paper, shows the westward motion of some of the lower spherical harmonics between 1830 and the present (by comparison, the axis of the main dipole has moved westward by only 8° during the same period). Within the errors of observation the angular velocity of the westward drift seems to be independent of latitude.

The westward drift appears closely associated with another remarkable geophysical fact, the irregularities in the observed rate of the earth's rotation.[14] These are obtained from the observations after the well-known regular slowing down of the earth's angular velocity attributed to tidal friction has been eliminated and the result compared with a uniform clock defined by planetary motion (absolute atomic clocks are now available and can be used for future observations). To give a quantitative idea of the change in the earth's angular velocity we may take the period of about 1890–1900 where there was a somewhat more than average change in ω amounting to $d\omega/dt = 1.5 \cdot 10^{-20}$ sec^{-2}; this represents a cumulative deviation of the earth clock by 10 seconds in 10 years. It has been demonstrated[15] that these fluctuations in all likelihood cannot be ascribed to any processes going on in the earth's crust or mantle, or in the ocean or atmosphere. A more likely explanation lies in a variable coupling between the mantle and core, plausible in view of the fact that the geomagnetic westward drift indicates the existence of a stationary torque acting between the two. The presumption has been further corroborated by Vestine[16] who succeeded in obtaining a curve of the variation in the rate of the westward drift, from 1900 on where the geomagnetic data are sufficiently accurate. The curve runs closely parallel to the curve of the irregularities in rotation, but has opposite sign throughout, a feature that will be explained presently.

Proceeding now to an interpretation of the westward drift and of the irregularities in the earth's rotation, we remark first that the observations mean that the mantle rotates faster than the top layers of the core where the secular variation originates (to be specific we may take a mean depth of, say 50 km, for this layer). Now in a rotating fluid mass in which there is turbulent exchange the angular velocity must necessarily decrease as we go from the center outward (a fact already discussed in Part I), and furthermore, the mantle must, in the stationary state, have the same angular velocity as the very top layer of the core. On purely mechanical grounds, then, any nonuniform rotation should have the opposite sign from that observed. The observed westward drift can only be explained on the assumption that a couple due to the ponderomotive magnetic forces acts between the mantle and the core. Similarly, the negative sign of the correlation between the irregularities of rotation and the irregularities of

---

[13] Bullard, Freedman, Gellman, and Nixon, Trans. Roy. Soc. (London) **243**, 67 (1950).

[14] D. Brouwer, Proc. Natl. Acad. Sci. U. S. **38**, 1 (1951).

[15] W. Munk and R. Revelle, Monthly Notices Roy. Astron. Soc., Geophys. Suppl. **6**, 331 (1952).

[16] E. H. Vestine, J. Geophys. Research **58**, 127 (1953).

the westward drift cannot be explained by mechanical coupling but again requires a time-dependent magnetic torque between mantle and core. All this implies that the mantle must have some degree of electrical conductivity, as a ponderomotive magnetic force cannot exist in an insulator. Postponing the discussion of the mantle's conductivity for a moment, we see from formula (I,4) that the mechanical force of the magnetic field is the cross product of $\nabla\times\mathbf{B}$ into $\mathbf{B}$. It turns out that $\nabla\times\mathbf{B}$ is essentially caused by the radial decrease of the toroidal field outwards, whereas the second factor, $\mathbf{B}$, is here the radial component of the poloidal field. Their cross product is a vector directed along circles of latitude (φ direction). On integrating over the mantle one obtains a net magnetic torque on the latter. This is essentially a unipolar induction effect, the mantle being, as it were, the armature of a unipolar induction motor. In the stationary state the differential torque between mantle and core must be balanced by an opposite frictional torque. This friction need not be mechanical; it may be shown that localized magnetic eddies will have the same effect. Furthermore, in the case of the irregular fluctuations of the torque, it must be assumed that any forces acting on the top layers of the core are rapidly communicated to the lower layers by means of the strong magnetic forces prevailing in the core; otherwise not enough moment of inertia would be available to achieve an effective transfer of angular momentum to and from the mantle.

Bullard[13] has given a schematical model of the westward drift which consists of three solid concentric metallic spherical shells in contact with each other, but able to rotate relative to each other. The inner two are highly conducting and represent the core; they are driven relative to each other so that the innermost one rotates faster, the whole being a crude model of the dynamo (Part I). The outermost sphere is weakly conducting and represents the mantle. It may then be shown that it tends to rotate faster than the intermediate sphere, corresponding to the westward drift. Elsasser and Takeuchi[17] have concentrated their attention on the hydromagnetic properties of the top layers of the core. They show that fluctuations in the toroidal component of the core's field, of the order of a few tenths of a gauss, which pass as transients into the mantle, are adequate to explain the observed irregularities of rotation.

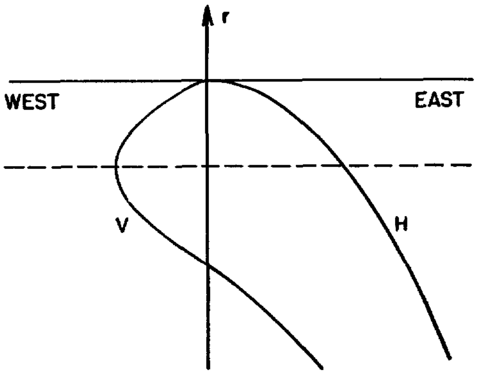


FIG. 3. Variation of zonal velocity and toroidal field with depth in the core.

If the lower surface of the mantle is "rough" (there being some modest mountains and valleys which might have been carved out by differential chemical action between mantle and core) the velocity of the fluid at the boundary must be essentially that of the mantle itself. On the other hand, at greater depth in the core the angular velocity must increase as we go downwards. To satisfy the observations of the westward drift and in agreement with theory[17] the velocity and toroidal field must then vary with depth in the upper parts of the core as shown in Fig. 3 (the relatively weak toroidal field penetrating into the mantle has here been ignored and the velocity is shown relative to the mantle).

The problem of the electrical conductivity (or rather, conductivity distribution with depth) in the mantle is clearly of much geophysical import. The first attack on this problem resulted from the extensive work of Chapman and his school on the ionospheric magnetic effects having periods of the order of a day. It was found that the electric currents that were assumed above the earth in order to represent these time-dependent magnetic fields gave rise to "image-currents" inside the earth. Detailed analysis[18] showed that these effects were compatible with a very low

---

[17] W. Elsasser and H. Takeuchi, Trans. Am. Geophys. Union **36**, 584 (1955); H. Takeuchi, and W. Elsasser, J. Phys. Earth **2**, 39 (1954).

[18] B. Lahiri and A. Price. Trans. Roy. Soc. (London) **237**, 509 (1939).

conductivity down to a depth of about 700 km, but from there on the conductivity must rise rapidly to at least 10^{-2} ohm^{-1} cm^{-1} at 900 km. The value of the conductivity deeper down, in the lower mantle, must satisfy two somewhat conflicting requirements: on the one hand it must be large enough so that the variable fields producing the irregularities in the angular velocity of the mantle can generate an adequate mechanical force there. On the other hand, the magnetic records (Fig. 1) show near-discontinuities of the second time derivative, meaning that the spectrum is freely transmitted through the mantle down to periods of only a few years. Quantitative studies[19] indicate that these conditions can be more readily fulfilled if the conductivity is assumed to increase very steeply with increasing depth, an assumption plausible on physical grounds. Runcorn[20] estimates the conductivity in the lower part of the mantle as of the order of 1 ohm^{-1} cm^{-1} which is to be compared with a value of several thousand ohm^{-1} cm^{-1} for the iron core.[1] The study of the problem is only in its beginning and promises interesting information about the physics of the mantle. The nature of the conduction process is either semiconductivity which has been experimentally demonstrated[20,21] in the laboratory for such rocks as olivine, considered representative of the composition of the mantle. Or else it might be ionic conduction which would be preponderant in rocks at higher temperatures. The mobility of ions, however, might be severely restricted by the high pressures, of the order of 10^6 atmospheres, prevailing in the lower part of the mantle, but this has not been proved.

FOSSIL MAGNETISM

The study of fossil magnetism deals with the behavior of the earth's field during earlier times as it has been preserved in rocks, clays, and also in human artifacts such as bricks. There are two principal types of magnetization of such bodies, namely, (1) cooling of a permanently magnetizable substance from a temperature above the Curie point to ordinary temperatures while in the earth's magnetic field, (2) the deposition during sedimentation (say at a lake bottom) of magnetized particles that are oriented by the earth's field. This process is usually complex because, if the particles are not spherical but, say disk- or rod-shaped, their geometrical shape will also influence their orientation on deposition. The subject of fossil magnetism has many ramifications into petrography, mineralogy, and crystallography; here we must confine ourselves to some brief remarks about the essentials in so far as they help us to evaluate the information that we may gather about the past history of the geomagnetic field.

Magnetization of rocks early aroused the curiosity of petrographers and geologists and from about the turn of the century the literature assumes appreciable proportions. The phenomena were, however, so complex that analysis of the findings in view of empirical regularity yielded little information of value. This changed only when it became possible to interpret fossil magnetism in detail in terms of the physical concepts of lattice structure and magnetic domains. The unraveling of the complex relationships is due in a large part to the theories of Néel[22] and to the work of Nagata[23] and his associates. It appears that the remanent magnetism of rocks, etc., constitutes a phenomenon rather different from conventional ferromagnetism, as in a piece of steel. In the first place, the permanently magnetized components of a rock constitute only a small fraction of its total mass and are almost invariably suspended in the form of fine grains. Although these grains might be very strongly magnetized, the remanent magnetization of the rock as a whole is as a rule extremely small. In the second place the magnetized components, being mostly oxides of iron ranging from simple Fe_2O_3 to more complex oxides containing other metals such as Ti and Mg,—these oxides are not really ferromagnetic but, as Néel[22] emphasizes, are antiferromagnetic.

The detailed investigation into the chemical and physical nature of these magnetized grains would not be necessary, were it not in order to

---

[19] K. McDonald, thesis, Utah, 1956; (to be published).

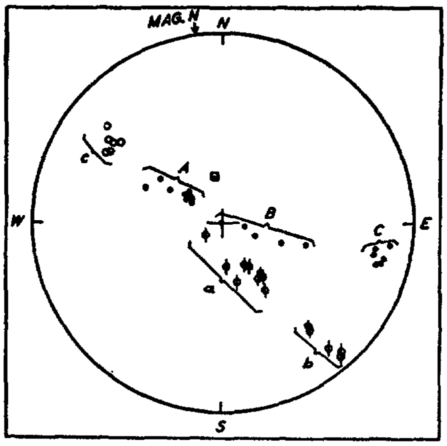
[20] S. K. Runcorn, Trans. Am. Geophys. Union **36**, 191 (1955); S. Runcorn and D. Tozer, Ann. géophys. **11**, 98 (1955).

[21] H. P. Coster, Monthly Notices Roy. Astron. Soc., Geophys. Suppl. **5**, 194 (1947).

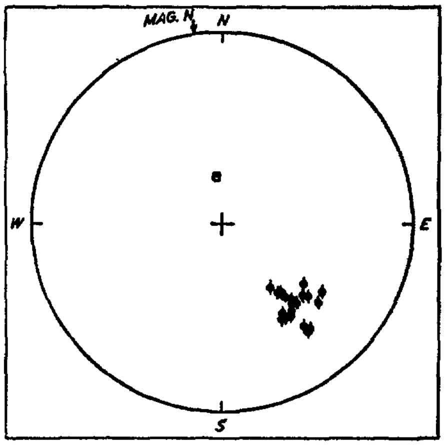
[22] L. Neel, Advances in Phys. (Phil. Mag. Suppl.) **4**, 191 (1955); see also L. Néel, Ann. géophys. **7**, 90 (1951).

[23] T. Nagata, *Rock Magnetism* (Maruzen Co., Tokyo, 1953).

gain some understanding of the magnetic *stability* of these bodies or, conversely, their lack of stability. Matters here would be quite simple if all rocks were either magnetically stable or else magnetically unstable, but unfortunately some rocks are found to be stable, representing the earth's field as it prevailed at their formation, while others are more or less unstable, having been remagnetized since. We cannot here go into the physico-chemical conditions for stability which are often not yet fully understood; it is clear, that in view of the very long time intervals involved in fossil magnetism, the results of laboratory measurements or of theory are necessary but by no means sufficient. We must, in addition, have recourse to stability criteria that refer only to the rocks as they are, and by their internal consistency prove beyond reasonable doubt that rocks have been magnetically stable through ages of geological length. Fortunately, tests of this type exist; the most significant ones were devised by Graham.[24] His technique consists in studying a bed of rock which has been folded after its formation. It is usually possible to determine quantitatively the local orientation in space of the folded strata. If now the remanent magnetization is determined at a number of locations having various slopes of the strata, one obtains in general widely scattered directions of the observed vectors of magnetization. (If on the other hand the observed vectors were all roughly in the direction of the present-day field, the rock would clearly be unstable and would have been remagnetized.) Now assume that each observed magnetic vector is corrected for the local inclination of the stratum so as to re-establish its direction relative to the original, geologically undisturbed bed. If in this process the original, widely scattered vectors become more nearly parallel, clustering about some mean direction, then the rock can be considered as stable and the mean direction of the cluster can be taken as the direction of the earth's field at the time of the formation of the rock. This is illustrated by Fig. 4 taken from Graham.[24] (The mode of representation is a stereographic projection of the hemispheres upon the equatorial plane.) There is also a converse to this stability test. Assume an extensive stratum shows fairly uniform remanent magnetization. If now by erosion, pebbles of this rock have been formed and have been embedded into a sediment, then the distribution of the magnetization vectors in the pebbles should be in random directions. If the formation of the sediment occurred not too long after the formation of the original material, the randomness of the magnetization in the pebbles furnishes an



OBSERVED POLARIZATIONS AND BEDDING POLES IN FOLDED ROSE HILL (SILURIAN) FORMATION; PINTO, MARYLAND, NOVEMBER 23, 1948
•=BEDDING POLES
♦=SOUTH SEEKING POLARIZATIONS
○=NORTH SEEKING POLARIZATIONS
■=EARTH'S PRESENT FIELD



POLARIZATIONS OF FOLDED ROSE HILL (SILURIAN) FORMATION CORRECTED FOR TILT OF BEDS; PINTO, MARYLAND, NOVEMBER 23, 1948
♦=SOUTH SEEKING POLARIZATIONS (CORRECTED)
■=EARTH'S PRESENT FIELD

FIG. 4. Showing magnetic stability. Top: magnetic vectors in space. Bottom: relative to bedding plane. (After Graham.)

---

[24] J. W. Graham, J. Geophys. Research 54, 131 (1949).

excellent criterion for the magnetic stability of the material.

Numerous rocks have in recent years been investigated by methods such as those outlined, or variants of these procedures. Great complexities arise because apparently some rocks are magnetically stable, some are unstable, and some are partially stable. For the details the reader may consult Nagata's book[23] or the recent review of Runcorn.[25] We shall report some of the main results of these investigations so far as they inform us about the past of the geomagnetic field.

Data for the historic period come partly from historically documented lava flows (Italy, Japan) partly from human artifacts such as bricks that have been fired and cooled through the Curie point at some known date. Thellier[26] has studied many such specimens in France and adjacent regions; perhaps his prize exhibit comes from the walls of kilns excavated not so long ago in Carthage, which apparently have been fired up to the destruction of that city. The observed variations in the direction of the magnetic force are consistent with the view that the secular variation has been of the same general magnitude in the past as it is now. Moreover, it has been possible to advance the brick method in a few instances to the point of measuring the absolute value of the past field, not only its direction. In the Roman period the field strength appears to have been larger by a factor of 1.5 than it is now. This shows that the main dipole has decreased in the past, though at a much lower rate than the -5% per century derived from direct observations since the time of Gauss.[1]

(Any appreciable variation of the geomagnetic field changes the total amount of cosmic radiation reaching the earth and hence the rate of production of radiocarbon. The half-life of radiocarbon is nearly 6000 years and the well-known dating procedures[27] should be rather insensitive to changes of the field of appreciably shorter periods. A curve given by Libby[27] shows very good agreement between known historical dates and dates obtained by the radiocarbon method, for 5000 years back. This might be an indication that the present dipole of the earth is not too far from the average over a longer period in the past.)

Turning now to prehistory, a valuable body of data has been unearthed in expeditions of the Carnegie Institution[5] studying varves and clays, and especially the content of drilling cores taken from the bottom of the Atlantic near Greenland. From this work which we cannot quote in detail there emerge two facts of interest for our purpose: While there is considerable scatter in the data there is also enough consistency to support the contention that the secular variation as we know it has existed for tens of thousands of years at least. Furthermore it appears that the irregular variations of the field are centered about true geographic North rather than about the present magnetic dipole axis. On the whole, the inclination of $11\frac{1}{2}°$ of the dipole axis to the geographical axis seems exceptionally large and has perhaps been less than half this value in the average over longer periods of the past. The idea that the time average of the geomagnetic field is a dipole centered on the earth's axis has been corroborated by statistical results derived from an extensive sampling[28] of flat-lying sediments in the northwestern United States (Fig. 5). The great majority of these rocks are of Tertiary origin and are between 5 and 50 million years old; only a few are older. Stability tests as described in the foregoing were not applied to these specimens. It is very unlikely, however, that the "permanent" magnetization of such rocks can follow the migrations of the dipole axis which has periods of the order of a few thousand years at the most; it is much more probable that even for unstable rocks the observed direction of magnetization represents an average over much longer periods. The declinations of Fig. 5 are seen to center closely on true North; the most probable inclination is about 6° smaller than the mean present-day inclination at these sites, corresponding to the fact that the sites are closer to the present-day magnetic pole, in Labrador, than to the geographical pole.

---

[25] S. K. Runcorn, Advances in Phys. (Phil. Mag. Suppl.) **4**, 244 (1955).

[26] E. and O. Thellier, Compt. rend. **212**, 281 (1941); **214**, 382 (1942); **222**, 905 (1946); **233**, 1476 (1951).

[27] W. F. Libby, *Radiocarbon Dating* (Chicago University Press, Chicago, 1955), second edition. I am indebted to Dr. J. Winckler and Dr. E. Ney of the University of Minnesota for pointing out to me the critical importance of the stability of the geomagnetic dipole in this respect.

[28] Torreson, Murphy, and Graham, J. Geophys. Research **54**, 111 (1949).

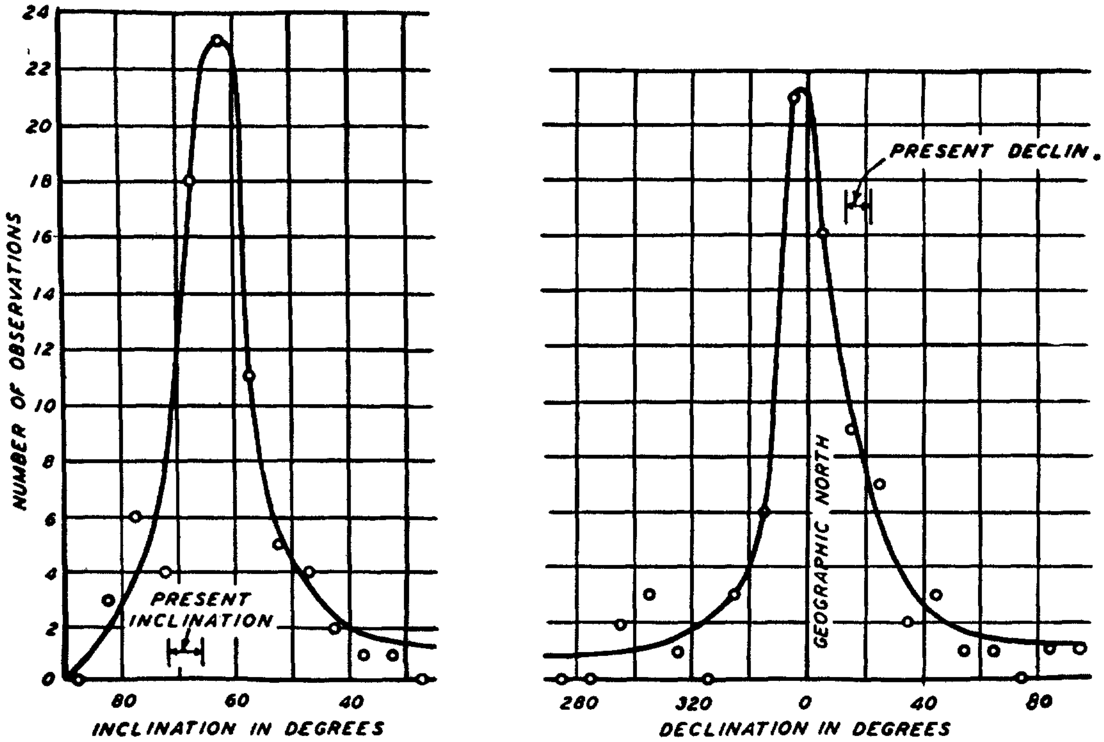


FIG. 5. Statistics of fossil inclination and declination (mostly Mesozoic). (After Torreson, Murphy, and Graham.)

On the dynamo theory one would expect that over sufficiently long times such an agreement between the geomagnetic and the geographical axis prevails. The fact that the two axes can deviate as much as is observed indicates that the hypothesis of numerous feedback loops used in Part I is not too good an approximation and that, in fact, the number of feedback eddies is relatively modest. Another indication that the feedback mechanism is none too homogeneous is furnished by the variation of the absolute intensity of the dipole field with time.

It has long been realized that the investigation of rock magnetism might furnish a means of ascertaining the truth of the hypothesis of polar wandering[29] repeatedly proposed by geologists. The idea arose out of the finding of paleoclimatology, especially the numerous remnants of paleozoic glaciations in Africa and the Pacific area. The pole was assumed to have been slightly north of Hawaii in the Carboniferous and to have migrated gradually in a somewhat meandering manner to its present position.[30] Such a motion requires a deformation of the ellipsoidal crust leading to shear stresses that can be calculated[31] and are at the least not in disagreement with the main tectonic shears actually observed. Ideas of this type call for a dynamical foundation. At present we do not have a dynamical theory of pole migration, but some valuable general arguments have been given by Gold.[32] Let us first define pole migration as a slipping of the solid crust and perhaps of the upper part of the mantle relative to the lower parts of the earth, assuming that the relative configuration of the crust remains unchanged.[29] It is most unlikely that such a process can take place in a conservative (reversible) fashion (after the example of the falling cat that lands on its feet). Motions in the core, for instance, normal to the earth's axis, even

---

[29] To be distinguished clearly from the hypothesis of so-called "continental drift" which so far lacks a foundation in the mechanics of solids and is commonly rejected by geologists.

[30] B. Gutenberg, editor, *Internal Constitution of the Earth* (Dover Publications New York, 1951), see Chap. 9.

[31] F. A. Vening Meinesz, Trans. Am. Geophys. Union 28, 1 (1947).

[32] T. Gold, Nature, 175, 526 (1955); see also a forthcoming paper of this author in the Monthly Notices, Roy. Astron. Soc.

extending over long times, would not be able to supply the energy required for the plastic deformation of the ellipsoidal crust. What Gold has in mind is an irreversible effect expressed by the frictional terms in the equations of motion. To illustrate the situation, assume that by tectonic forces in the upper strata of the earth a hump is created someplace. In a rigid gyroscope this would merely give rise to the familiar precession of the asymmetrical top. Now it is clear that in an internally deformable system the only configuration of long range stability is one where the hump has moved to the equator. Hence forces will be set up that tend to achieve this, and they may slowly produce the requisite plastic slippage over the lower layers. Gold also points out that situations can arise where the pole becomes "trapped" and cannot move far from a given position without a major perturbation. This he claims applies to the present position of the pole. It is certainly to be hoped that a mathematical theory of slow, irreversible deformations of a solid gyroscope will be developed in the not too distant future.

This is the more desirable since we should soon have definite observational answers to the problem of pole migration. A review of the present status of geomagnetic data has been given by Runcorn,[25] who first systematically tackled this problem. He locates the geographical North pole somewhat south of Hawaii in Precambrian times; in the Silurian it is already in Northern Japan and in the Eocene just off the northern coast of Western Siberia, not too far from its present position. This rather early time of migration reflects the fact that the magnetic data from rocks of more recent geological periods tend to show a field direction that agrees with the present one. The basic data for Runcorn's attempt come from paleozoic rocks in England; since, however, pole migration produces a world-wide magnetic variation, we can hope to see these findings checked by paleomagnetic studies on a much broader scale than have been possible in the early stages of this subject. New data have now been obtained by Graham[33] who showed that measurements in paleozoic rocks, mostly in New Mexico and Arizona, give locations of the pole which agree, within the errors of measurement, with the English data. These results have recently been corroborated by Runcorn[34] from rocks of similar locations.

We finally come to one of the most intriguing findings of paleomagnetism, namely, the strong suggestion, if not as yet complete certainty, that the earth's field has frequently *reversed* its direction in the geological past. This is not a more rapid pole migration, but an outright "flopping over," by 180° effectively instantaneous on the geological scale. If such reversals are due to modifications or fluctuations of the dynamo mechanism we might indeed expect them to occur at rates characteristic of the processes involved, say within a few thousand years at the most. Any such interval is negligible from the geological viewpoint. Reverse remanent magnetizations of lavas and other rock formations have long been noted, but a phenomenon of an almost ludicrous character was then discovered which seriously complicated the investigations. This is the spontaneous reversal of the direction of remanent magnetization rocks. It was only after this phenomenon was thoroughly clarified that one could, fairly recently, put some trust into the evidence for actual and in fact relatively numerous reversals of the earth's field. Surprisingly enough, there is not just one mechanism of spontaneous reversal of rock magnetization, but several, at least four, which have been thoroughly discussed by Neel[22] to whom we owe the theory of such phenomena. This article is not the place to discuss these intriguing peculiarities of solid-state physics, but in order to satisfy the reader we shall crudely indicate a couple of these processes: Suppose in a complex crystal there are two kinds of ion or atoms, A and B such that the A's have their spins parallel by ferromagnetic coupling, whereas A and B are antiparallel by antiferromagnetic coupling. If the magnetic moment of B is larger than that of A, the net magnetization is in the direction of B. Assume now that by a diffusion process B is gradually replaced by an ion, C, whose net spin is subject to the same couplings as B but whose magnetic moment is smaller, and is in fact smaller than A. The effect of this substitution will be a reversal of the domains of magnetization of the crystal. Another mechanism is as follows: Suppose the

---

[33] J. W. Graham, J. Geophys. Research **60**, 329 (1955).

[34] S. K. Runcorn, Nature, **176**, 505 (1955).

grains of two ferromagnetic substances, A and B tend to grow close together; they will then be oriented so as to take up each other's flux, that is, they will be arranged as two magnets antiparallel and next to each other. Now assume that A is in general stronger than B so that the net magnetization of the specimen is in the direction of A. If now A is slowly leached out by chemical action, the grains B remain and the magnetization is reversed. The general principle of these reversals is clearly that the orientation of the spins in a magnetic domain can survive chemical actions in the crystal itself or in its neighborhood. Neel has not only discussed these mechanisms at great length, but in some cases has been able to demonstrate their existence experimentally. Moreover, by processes closely related to those just mentioned, a rock can become reversely magnetized on being cooled down from a temperature above the Curie point. Nagata[23] has found an actual rock with this property and has analyzed it extensively.

Basically this spontaneous reversal of rock magnetism seems something of a freak phenomenon; nevertheless it can be quite serious in field explorations of rock magnetism. Thus in aerial magnetic surveys of the Adirondack mountains by the U. S. Geological Survey numerous magnetic anomalies were found. It was then shown[35] by an analysis of actual samples that there is a distinct difference in the average chemical composition of the reversely and the normally magnetized specimens. Other, remarkable reversals of Néel types have been found in Japan.[36] In view of such experiences it is hardly necessary to say that the investigators who have found evidence for actual reversals of the earth's magnetic field have taken every precaution to ascertain that these reversals are not of the spontaneous, chemical type. Frequently the geologist finds a number of lava flows piled on top of each other in regions which were once volcanically active. If there are reversals of the remanent magnetization of some layers with sharp magnetic boundaries but absolutely no discernible chemical change at the same boundary, the likelihood of an actual reversal of the earth's field becomes large. In recent years, a number of such cases have been carefully investigated.[25] There is, furthermore, now extensive evidence[37,38] that where reversely magnetized lava flows are spread over pre-existing rocks or clays, the layer of the rock immediately below the lava acquires the same magnetization as the latter, independently of the particular petrographic characteristics of the material. This is no doubt due to the material's being heated above the Curie point of its magnetic components by contact with the lava. As a result of such cumulative evidence many investigators have now become convinced that reversals of the earth's field in the geological past have occurred. Figures 6 and 7 (from Runcorn[25]) give examples of actual data. In each case solid circles indicate one, open circles the opposite direction of the field. Figure 6 represents data from Columbia River basalt flows (Miocene) whereas Fig. 7 comes from Torridonian Sandstones in Scotland (Late Precambrian). The selective preference of nearly opposite magnetizations over a random distribution of the vectors is clearly evident. Comparison of the two figures may also serve as an illustration of pole migration in the intervening period (see also Fig. 4).

For a more detailed review of the observations Runcorn's paper[25] should be consulted where references may be found. We solely mention here Hospers' extensive data on multiple reversals in Icelandic lavas; a great deal of work done on Japanese lavas by Kawai and others; studies by Roche on reversely magnetized lavas in France and by Hospers in Ireland. Data recently obtained by Runcorn on lavas in Arizona add further corroboration. Most of the lavas investigated are of Tertiary origin. The Pilansberg dikes of South Africa, a vast system of volcanic intrusions, of paleozoic age or older, show mostly reverse magnetization which has been studied by Gough. Sedimentary rocks withstood attempts at interpretation longer than lavas, but recently Clegg, Almond, and Stubbs have succeeded in getting consistent and significant results from English sediments. In their data, opposite directions of magnetization seem to occur with about equal probabilities.

Runcorn estimates that there have been

---

[35] A. Buddington and J. Balsley, J. Geomag. Geoelec. **6**, 176 (1954).

[36] E. Asami, J. Geomag. Geoelec. **6**, 145 (1954).

[37] A. Roche, Compt. rend. **236**, 107 (1953).

[38] Einarsson and Sigurgeirsson, Nature **175**, 892 (1955).

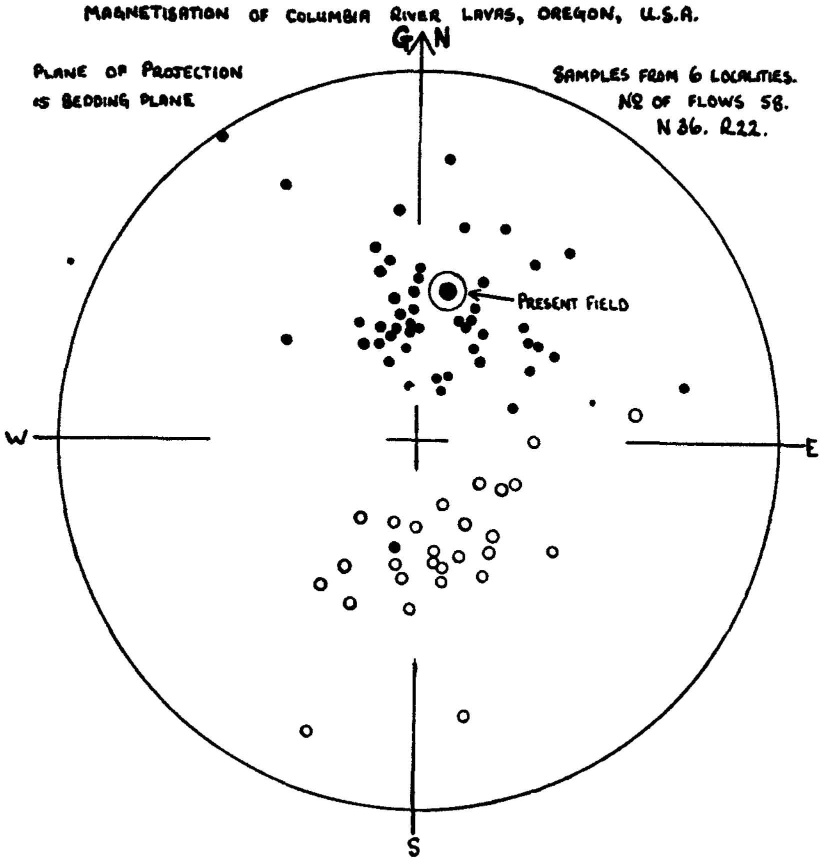


FIG. 6. Showing near-equal occurrences of direct and reverse remanent magnetizations in lava flows of Columbia River basin. (After Campbell.)

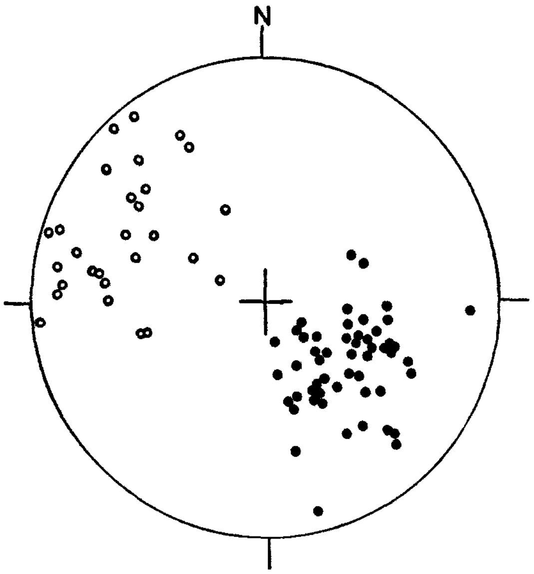


FIG. 7. Showing near-equal occurrence of direct and reverse remanent magnetization in paleozoic sandstones of Scotland. (After Irving.)

several hundred reversals of the earth's field during geological history. There are also some indications that opposite polarities may have occurred with equal probabilities, but the accumulation of far more data must be awaited before clearcut conclusions can be drawn.

No dynamical theory of the reversals has as yet been given. A little reflection will show that with some imagination one can construct not one but several models of reversals, but we shall not try to enter into details. It is noteworthy, however, that by the evidence of observation the dynamo problem does have two types of solutions, the mean-stationary type as in the earth and the migratory (dynamo-wave) type as in the sun. It is therefore not impossible that a modification of the convective regime (such as a transient intensification) will lead to a condition where a reverse toroidal field is created at some depth in the core, which then grows and migrates outward by a process similar to the migration of the solar dynamo.

SOLAR AND STELLAR MAGNETISM

The direct evidence for the existence of magnetic fields on the sun and stars comes from spectroscopic observations. Solar and stellar spectral lines are as a rule intrinsically broad, partly owing to turbulence in the layers where they are produced, and partly on account of the Doppler broadening that arises from rapid rotation (especially in "early" type stars). An actual splitting of the lines under Zeeman effect is only observed in magnetic fields over 1000 gauss such as prevail in sunspots. For weaker fields one must be satisfied with lines that show an excess broadening due to the magnetic field, detectable in terms of a partial polarization of the line wings. Usually, longitudinal Zeeman effect is studied, corresponding to the magnetic field component directed along the line of sight. In "normal" Zeeman splitting this converts a line into a doublet whose components are circularly polarized in opposite senses. When the line is not actually split a partial circular polarization of the line edges can still be determined by means of suitable analyzers. Numerous lines of the solar spectrum show Zeeman effect but a few of these, especially iron and chromium lines, are especially suited because they are comparatively sharp and have relatively simple and wide Zeeman patterns.

In the two decades following Hale's discovery of sunspot magnetism (1908) magnetic fields on the sun were extensively studied at the Mount Wilson Observatory, and most of our detailed knowledge comes from there. The magnetic resolving power of these measurements is about 50 gauss. More recently, Babcock[39] has designed a photoelectric scanning device for solar magnetic fields which has been in operation since 1952. This machine is fully automatic; it continuously compares the polarizations of the two edges of a line while scanning the surface of the sun. It has remarkable optical properties (resolving power of the grating used is 600 000, dispersion achieved is 11 mm/A) to the point where further increase of resolving power would be useless on account of the natural line width. Figure 11 shows some typical scans, the field strength being given by the vertical deflection above or below the base line; the distance between adjacent base lines represents very nearly 1 gauss.

The measurement of the magnetic field of stars was successfully achieved some years ago by Babcock,[40] and a sizeable number of magnetic stars have been studied. The optical technique is essentially similar to the one used on the sun; the equipment is attached to the 100-inch telescope, sometimes to the 200-inch, and the spectrum is photographed. The accuracy of these delicate measurements cannot of course compare with solar observations; it depends greatly on the nature of the stellar spectrum and the diffuseness of the lines; the probable errors indicated are usually of the order of several hundred gauss.

A few words about the constitution of the sun as a whole are conducive to a better understanding of solar magnetism. In the first place it must be realized that the layer into which we can see, the photosphere, provides a rather sharp cutoff in depth for direct investigations. The scale height of the solar atmosphere (the height over which the density changes by a factor of *e*) is only about 100 km in the region of the photosphere.[41] Clearly then, it does not take very many hundreds of kilometers to go from a region so rarefied as to be practically transparent to one so dense as to be practically opaque. The photosphere is thus very thin compared to the solar radius of $7 \cdot 10^5$ km. We are in the position of a man who looks at an object with a microscope and can actually observe only a very thin slice of it.

A matter of prime importance in solar and stellar hydromagnetism is the stability of stratification in the various layers of the star. If a layer is stably stratified, motion, if present at all, can only take place along the equipotential surfaces or very nearly so. It may be shown on general grounds that when motions are essentially two-dimensional the lines of force cannot be twisted in such a way as to produce unlimited amplification (see Part I). Hence dynamo mechanisms can be maintained only in regions where there is three-dimensional motion. This means in practice that active thermal convection resulting from unstable stratification must exist. The well-known stellar model of Eddington in which radiative transfer is the controlling agency in the interior

---

[39] H. W. Babcock, Astrophys. J. 118, 387 (1953).

[40] H. Babcock and T. Cowling, Monthly Notices Roy. Astron. Soc. 113, 357 (1953).

[41] G. P. Kuiper, editor, *The Sun* (University of Chicago Press, Chicago, 1953).

of the stars leads to highly stable stratification with an extreme reduction of all possible motion. When the carbon cycle was introduced as the source of stellar energy, its pronounced temperature sensitivity led to the conclusion that regions in which this process operates must be convectively unstable. For some time it was then believed that the sun has a convective core but was otherwise highly stable. Since magnetic fields could not possibly be generated in the quiescent region, a great deal of ingenuity went into the construction of models[42] that allowed the transport of magnetic fields from the solar core where they were thought to originate, through the stable layers to the surface. In recent years our ideas about the internal constitution of the sun have changed rather radically. In the first place, the carbon cycle has been replaced by a proton-proton reaction as the main source of solar energy. This process is much less temperature sensitive, and if a convective core exists at all, which is doubtful, its convective activity is likely to be weak. Secondly, it has been found that the outer layers of the sun are convectively unstable. This region is designated as the hydrogen convection zone. It is fairly easy to see that when the specific heat of a gas increases with increasing depth (increasing temperature) in a layer, then the layer is less stably stratified than a region of constant specific heat. (The reader not familiar with this kind of argument may look up the definition of stability of stratification as function of specific heat and temperature gradient given in any text on meterology.) The hydrogen convection zone occurs in the temperature interval where the specific heat of atomic hydrogen is doubled owing to its progressive ionization. Since hydrogen is, of course, the main constituent of stars, this has a profound influence on the layers involved. In the sun, the upper limit of the hydrogen convection zone is somewhere near the lower edge of the photosphere. Recent computations[43] put the total depth of this zone at between 0.65 and $1.65 \cdot 10^5$ km, or 10–25% of the solar radius. The atmospheric layers of the sun above the hydrogen convection zone are stably stratified. The so-called solar activity,[44] that is the deviation of the solar envelope from its average condition, which sometimes includes phenomena of extreme violence, appears more and more as a byproduct of the dynamics of the underlying convection zone. As our knowledge of these phenomena progresses it is becoming abundantly clear that magnetic fields play a fundamental role in the layers of the sun accessible to us. We must conclude that there are efficient dynamo processes in the convection zone, of the type outlined in Part I, or similar, and solar hydrodynamics appears far more as a matter of hydromagnetism than of ordinary fluid mechanics.

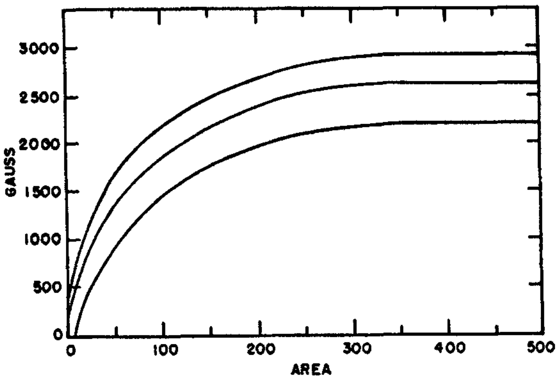


FIG. 8. Statistical distribution (mean and mean deviations) of magnetic field intensity (gauss) in sunspots. Abscissa is area of umbra in 10^{-6} of solar disk. (After Richardson.)

The magnetic fields of sunspots are easily the most conspicuous feature of stellar hydromagnetism. Summaries of the widely spread observations may be found in the book already quoted,[41] in Unsöld's book,[45] and in a monograph by Nicolet.[46] Let us begin by saying that *all* sunspots have magnetic fields, the field being as a rule the stronger the larger the spot. Figure 8, due to Richardson, shows the quantitative relationship between average area and average field in gauss (middle curve; the upper and lower curves indicate the statistical deviations from this mean). The abscissa gives the area of the "umbra" in units of 10^{-6} of the area of the solar hemisphere. (The umbra is the dark, central part

---

[42] See for instance, H. Alfvén, *Cosmical Electrodynamics* (Clarendon Press, Oxford, England, 1950).

[43] E. Vitense, Z. Astrophys. 32, 135 (1953).

[44] See the article on Solar Activity by K. O. Kiepenheuer in "The Sun," reference 37.

[45] A. Unsöld, *Physik der Sternatmosphären* (Verlag Julius Springer, Berlin, Germany, 1938).

[46] M. Nicolet, "Le Soleil," Inst. Roy. Meteorol. Belgique. Misc. Fasc. 11 (1943).

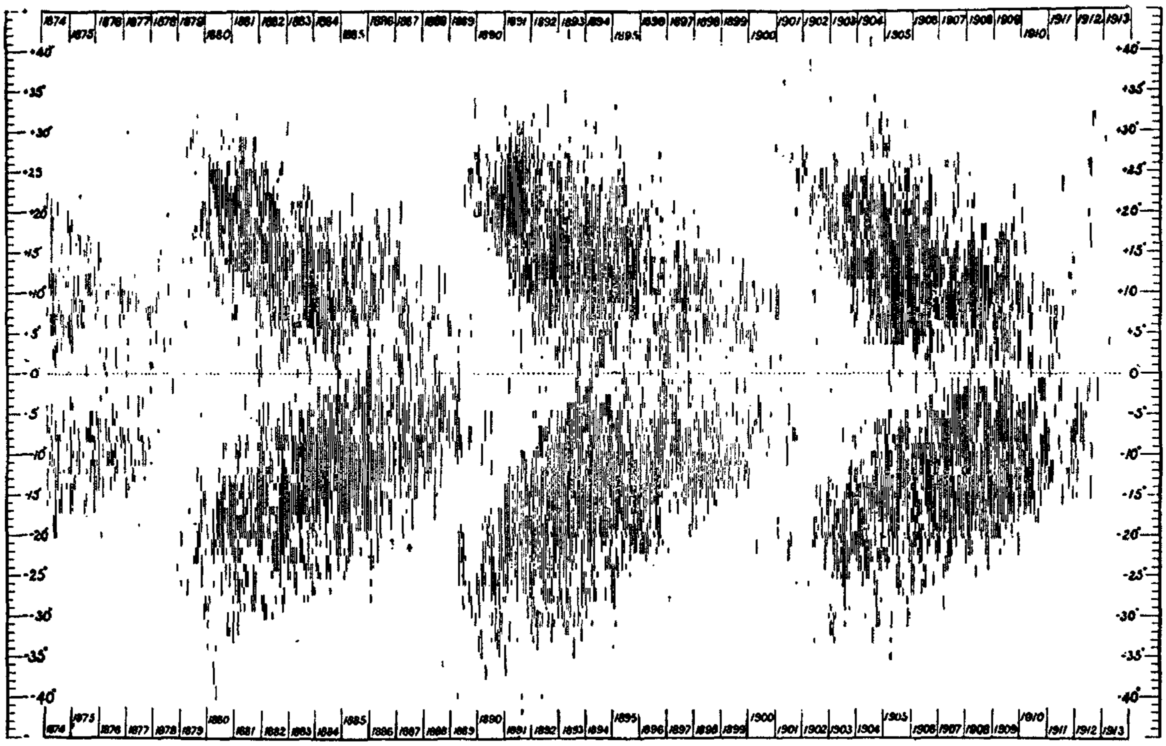


FIG. 9. "Butterfly" diagram showing variations in latitude of sunspots with the sunspot half-cycle. (After Maunders.)

of the spot, surrounded by a less dark edge, the penumbra.) We see from Fig. 8 that a spot of sufficiently large area has in the average a field strength of about 2600 gauss. Fields as high as 4500 gauss have been observed. The field direction is such that in the center of the spot the field vector is normal to the solar surface, but as one goes outwards from there the vectors become progressively tilted away from the normal. The following is the mean angle of inclination relative to the solar radius as function of the fractional distance from the center to the edge of the spot:

fractional distance	0	0.3	0.6	0.9
outward inclination	0°	18°	49°	73°.

A characteristic feature of sunspots is the motion of the fluid radially outwards, as detected by the Doppler displacements of lines in the spot (Evershed effect). The velocity is of the order of 1 km/sec and rises to a maximum around 2 km/sec at the edge of the penumbra.[44] We must remember that a sunspot is a highly anisotropic structure: to see this it is only necessary to compare the scale-height of 100 km in the photosphere with the lateral extension of a spot, a larger spot having a diameter of several times 10^5 km. On the other hand the velocity of outflow is quite fast and the material of the spot is therefore renewed in short order. It follows that diffusion, even eddy diffusion, is not likely to alter the field in the spot radically during the time it is convected through the spot by the flowing matter. The observed slant of the magnetic field lines must then be a phenomenon in the photospheric layer; no doubt a flux tube must reach from the spot into the deeper layers of the convection zone. We thus infer a mushroom-shaped figure for the field distribution in a spot. Matter must flow along the field lines into the spot to replenish the field of the latter, as the lifetime of a spot is quite commonly of the order of several weeks.

The ponderomotive forces of the magnetic field are by far the most important phenomenon for the dynamics of a sunspot.[47] To estimate them, consider a cylindrical tube with a homoge-

---

[47] L. Gurevich and A. Lebedinsky, J. Phys. Sowjet Union. **10**, 327, 425 (1946).

neous field B in it and zero field at the outside. The "magnetic pressure" which this field exerts on the surface of the cylinder is $B^2/2\mu$ or, to use unrationalized cgs units for a moment, $B^2/8\pi$ which at once gives an expansive pressure of about $3\cdot10^5$ dyn/cm^2 assuming the average field of a large spot (2600 gauss, Fig. 9). On the other hand, the hydrostatic gas pressure in the solar photosphere is of order 10^4–10^5 dyn/cm^3 and reaches $3\cdot10^5$ dyn/cm^2 only at the bottom of the visible layer.[41] It is therefore at first glance difficult to account for hydrostatic near-equilibrium in a sunspot such as seems required by the relatively long lifetime of spots. Now observations show that sunspots form depressions in the solar surface; if the depression amounts to several scale heights, as is generally assumed, the gas pressure at this lower level would be so large as to appreciably exceed the magnetic pressure.

A theory, even a purely static one, of sunspots offers great difficulties. A study of hydrostatic conditions in sunspots has been made by Parker.[48] As we go deeper into the convection zone, things get even more obscure. Since the flux-conservation theorem holds for a compressible fluid, the field would increase as $\rho^{\frac{2}{3}}$ provided the compression was isotropic. On such an assumption Gurevich and Lebedinsky[47] arrive at very large fields indeed, of the order of 10^5–10^6 gauss farther down in the convection zone. Parker[48] thinks that the forces produced by the field will render the compression or expansion highly anisotropic so that such assumptions are not required. A great deal of theoretical analysis remains to be done. The "coolness" of a spot (about 1000° below solar surface temperature) is undoubtedly related to the hydrostatic balance of forces in which the gas pressure plus magnetic pressure on the inside must equal the outside gas pressure, but it seems difficult to account for the actual temperature in the spot without an elaborate model of spot dynamics. (Given the small scale height and pronounced anisotropy of the spot region, most surface phenomena are no doubt largely determined by dynamical processes in deeper layers, difficult of access.)

So far we have dealt with an individual spot. We now consider sunspots as a mass phenomenon, indicative of broader aspects of solar dynamics. Very commonly there appears a conglomeration of a number of smaller spots in place of one large spot. By far the most important grouping is the so-called *bipolar group* where there are two spots, or two spot groups, or a spot and a group, separated from each other by an appreciable distance, several times the diameter of a large spot. These bipolar spot groups are considerably more numerous than single spots, and there is every reason to believe that a single spot is merely a degenerate form of a bipolar group, the other half of the bipolar group taking the form of a widespread, relatively weak field (as revealed by Babcock's magnetograph, see below) rather than of a localized, strong field, as in a spot. The line connecting the two components of a bipolar group is very approximately parallel to a circle of heliographic latitude, that is the two components have nearly a heliographic east-west orientation relative to each other. The observed law of magnetic polarity is as follows: the western spot or spot group has a fixed magnetic polarity for all bipolar groups appearing during one 11½-year sunspot cycle, and the eastern spot or group has the opposite polarity. This polarity is reversed from one sunspot cycle to the next. (Therefore the 11½-year sunspot cycle should really be called a half-cycle. The law of polarity in the southern hemisphere is the exact reverse from that valid for the north hemisphere.

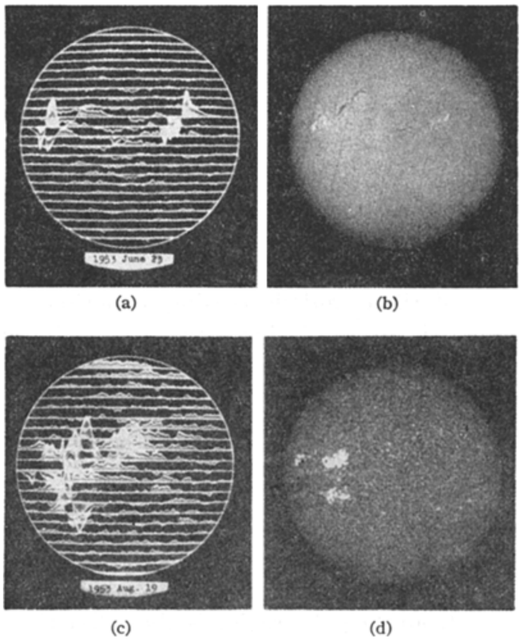
At the beginning of each sunspot half-cycle spots or spot groups appear in a solar latitude of about 30°. While the individual spot does not change its latitude appreciably during its lifetime, subsequent spots appear at progressively lower latitudes. This continues until toward the end of the 11½-year half-cycle the spots are only a few degrees from the equator. At the same time the first spots of the next half-cycle appear near 30° of latitude, but with reversed polarities. The structure of the sunspot cycle is best illustrated by Maunders' butterfly diagram[41,46] shown in Fig. 9. The abscissa is time, the diagram comprising 40 years, or 3½ half-cycles, the ordinate is solar latitude. Each vertical line corresponds to one solar rotation period (26 days) and indicates the scatter in latitude of spots observed during that period.

In view of the preponderant bipolarity of

---

[48] E. N. Parker, Astrophys. J. 121, 491 (1955).

sunspots and the reversal of polarity as between the two hemispheres, it seems almost certain that the spots owe their origin to strands of a solar toroidal field in the hydrogen convection zone which have been heaved up to the surface (perhaps by an effect of "magnetic buoyancy")[48] as shown in Fig. 10. The general pattern of this interpretation is old; such a scheme was proposed over 30 years ago by Bjerknes who, ignoring the magnetic effects, interpreted as a vortex filament what now is the toroidal field. It follows from the observations that this field migrates gradually, in the course of a half-cycle from middle latitudes towards the equator, and in the next period a toroidal field of reversed sign appears and wanders likewise to the equator. These observations have been represented in terms of the concepts and formalism of the dynamo theory by Parker's model of migratory dynamo waves quoted in Part I. However, we do not yet know at what depth in the convection zone these processes are active and what order of magnitude of field they can produce in these depths.



(a) (b)

(c) (d)

FIG. 11. (a) Magnetogram record of bipolar region related to intensified H_α emission shown in (b). Strong unipolar field in (c) related to intensified CaII emission shown in (d). Distance between traces = 1 gauss. (After Babcock.)

The older methods of solar magnetic analysis developed by Hale and his co-workers give us ample information about sunspots whose fields range from a few hundred to a few thousand gauss, depending on the size of the spot. These methods have not been so successful for the smaller fields outside of spots. New information has now become available,[49] through Babcock's solar magnetograph mentioned previously. This instrument averages over a sizeable fraction of the solar surface, as may be seen from an inspection of Fig. 11. Sunspots do not seriously affect the instrument since the length of the slit is much larger than the diameter of an average sunspot and since, moreover, the amplifier saturates at about 10 gauss. Babcock's observations show that there are frequent magnetic regions of considerable extent. The great majority of these regions are bipolar with the same signs for the magnetic polarity as indicated above for bipolar sunspot groups. On occasion, sunspots of the appropriate polarity are associated with these magnetic regions which are, however, of vastly larger area than any spot. In case of such association it has been found that often the bipolar region of weak but extended field outlives the associated spots. This does not mean that such magnetic regions are always very stable; often they change quite rapidly in structure and extent. They seem associated with other, spectroscopic features of solar activity, as may be seen from Fig. 11. One of the significant findings of Babcock is that one polarity of a bipolar sunspot group can be replaced by a weak, widespread field, the net magnetic flux issuing from the surface being then more nearly zero. This seems to corroborate our model, Fig. 10; it also might indicate that the toroidal field farther down in the convection zone is more uniform and is less intense than sunspot fields. The latter would then be secondary concentrations of flux occasioned somehow by the break-through into the outer solar layers.

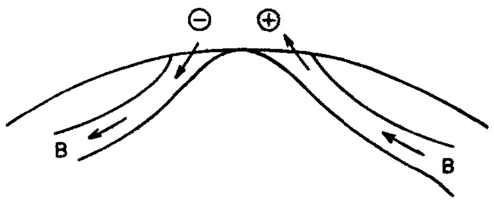


FIG. 10. Formation of a sunspot pair.

---

[49] H. W. Babcock and H. D. Babcock, Astrophys. J. 121, 349 (1955).

Babcock has also investigated the existence of an over-all solar field corresponding to a main dipole. Such a field seems to exist fairly definitely, but only in latitudes above 55°. Its average magnitude in the photosphere is 1–2 gauss, its sign relative to the rotation is the opposite of that of the earth's field. Furthermore, the effective flux seems to be subject to rapid and random fluctuations, increasing sometimes by as much as a factor of three and decreasing sometimes to zero. There is no correlation between such fluctuations in the northern and southern hemispheres. If we remember that what is actually observed is a field in a very thin, highly conducting region bounded on both sides by thick, highly conductive regions, and that in all these regions there prevail fluid motions carrying the field along, then most of these results seem slightly less surprising. The total average magnetic flux emerging from the polar region is about five times the flux going in and out of a typical bipolar region and is less than the flux emerging from a large sunspot. This again indicates that, contrary to the case of the earth, no coalescence of the solar field into a sun-wide main dipole is taking place; the solar hydromagnetic phenomena seem to be controlled in a more local fashion by processes going on in certain sections of the convection zone.

Leaving aside hydromagnetic phenomena in the higher layers of the sun which will be dealt with briefly later, we now turn to stellar magnetic fields. A star can show an observable magnetic field only when there is a strong line-of-sight component of one sign over most of the star's surface. If the field of the star is a dipole, the observed field strength is given by

$$B_{\mathrm{ob}} = (D/4)B_p \cos\lambda,$$

where D is a numerical factor taking account of limb darkening (deviations from Lambert's law); usually D is fairly close to unity. B_p is the field strength at the pole and λ is the inclination of the dipole axis to the line of sight. The observed fields (of the order of some thousands of gauss) are therefore smaller, as a rule, by an appreciable factor than their corresponding B_p. Furthermore, only rather bright stars can be spectroscopically analyzed to the extent necessary, and Babcock's stars are between the third and the seventh magnitude. In rapidly rotating stars the Doppler broadening of the lines is usually strong enough to obscure the magnetic effects unless the angle λ is relatively small. In his 1953 summary[40] Babcock tabulates 35 stars in which Zeeman effect has been measured in detail. In about 20 more stars magnetic fields are strongly suspected, but more measurements are required. Some 65 stars had to be rejected on account of too few spectral lines or too much rotational line broadening.

The conditions for the successful discovery of magnetic stars are clearly rather stringent. Since a large number of the stars investigated do actually show magnetic fields, Babcock concludes that magnetic activity must be widespread among stars. The actual distribution of the magnetic stars among spectral classes may be considered as a significant corollary of the dynamo theory: We have seen in Part I that the simultaneous presence of convection and rotation is prerequisite for dynamo action in our theory. Now "early"-type stars are found spectroscopically to rotate, as a rule, rapidly; "late"-type stars rotate slowly. There is evidence from the theory of stellar structure[50] that the earliest types, the *O* and *B* stars, do not have convective envelopes. These begin only with the *A* stars where they are relatively shallow; they grow deeper in later types. Of Babcock's 35 magnetic stars, 25 are of type *A* and some others of types *F* and *M*; some details of the observations seem to corroborate the view that in *A* stars the convective layer is rather thin.[51] On this view the sun, being a rather "late" star has a deep convective layer, but does not seem capable for some unknown reasons to develop an intense uniform field.

The fields of the measured magnetic stars range from a few hundred up to 10 000 gauss, the majority being of the order of 1000–5000 gauss. All objects sufficiently investigated were found to be magnetically variable. Some of these stars

---

[50] The author is greatly indebted to the astronomers, B. Strömgren, L. Spitzer, and L. Aller for detailed information on this subject. The absence of a convection zone in the early stars is due to the fact that they are so hot that hydrogen is already completely ionized at their surface. Quantitative studies of special stars of these types may be found in the literature, but no general survey seems to exist.

[51] H. W. Babcock, personal communication.

were known previously to be spectrum variables. The magnetic variation of most of the stars is irregular; some show periodic variations of nonuniform amplitude. In no case is the observed variation of the field simply a sinusoidal curve; in the simplest case a periodic field (containing some higher harmonics) is superposed upon a constant field (Fig. 12, upper curve). Often the curve is much more irregular (lower part of Fig. 12). The stars of Fig. 12 are also spectrum variables and some spectral characteristics are indicated in the diagrams.

The interpretation of magnetic variability offers many difficulties and is not yet far advanced. As Cowling[40] remarks, two immediate interpretations present themselves: The simplest model is that of the "patchy" rotator; one imagines that in such a star the phenomenon of sunspots occurs on a very large scale so that large but irregularly bounded fractions of the star's surface exhibit strong magnetic fields. If the axis of rotation is inclined relative to the line of sight the star will, in the course of its rotation, present different magnetic aspects. According to investigations by Deutsch[52] into the nature of the spectral changes, many of the findings on magnetic variables can be explained in this way.

The second type of model assumes that the magnetic variability is due to fluid motions inside the star. If, for instance, the star's photosphere undergoes a large scale oscillatory motion, the lines of force would be dragged along and so offer various aspects to the observer. Schwarzschild[53] proposes oscillations in which the fluid moves back and forth along the meridian. It is hard to see, however, how such simple oscillations can subsist in a rapidly rotating system. A related proposal of this author tries to explain reversal of polarity during the sunspot cycle as the result of torsional oscillations in the interior of the sun which would amplify the toroidal field in alternate senses during each half-cycle. Again, it is difficult to see how such oscillations can survive in the presence of a doubtlessly very high eddy viscosity in the convection zone.

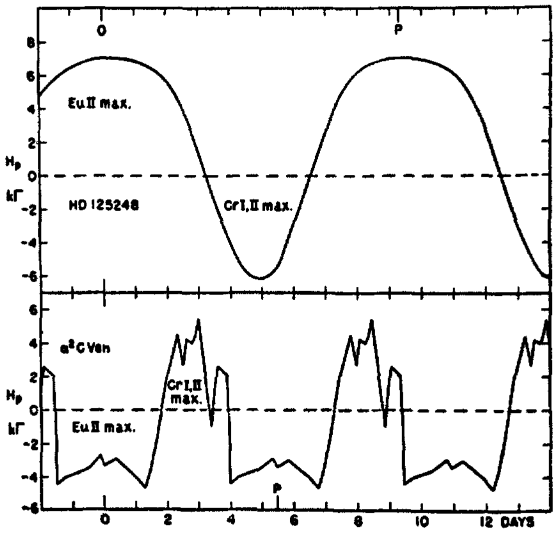


FIG. 12. Field-strength curves of two magnetic variables. (After Babcock.)

At the present time an interpretation of Babcock's results in terms solely of a rigid rotator model appears inadequate, at least in some cases,[51] and some sort of internal oscillation seems to be called for. Now any oscillation whatever corresponds to a periodic transformation of energy as between two forms. In Schwarzschild's models, just discussed, the two forms of energy are kinetic energy of the fluid and magnetic field energy, respectively, the whole being rather closely akin to the mechanism of Alfvén waves (Part I). In Parker's dynamo waves (Fig. 7, Part I) on the other hand, the energy exchange is between two mutually orthogonal types of modes of the magnetic field, poloidal and toroidal. It seems fairly clear, quite apart from such a specific model, that the sunspot cycle represents a 23-year oscillation, and there seems to be no valid reason to deny that oscillations of a similar dynamical nature might, under suitable conditions, have periods several hundred times shorter, as found in some stars. If the assumptions of the dynamo theory are correct, very strong stellar fields should be an indication of rapid rotation combined with strong convection, and the intense and complicated eddy motions which this would engender seem to be most unfavorable to sustained oscillations which involve a transfer between kinetic and magnetic energy. "Dynamo waves" involving an exchange of energy between two types of magnetic modes seem more likely, although at the present state of our

---

[52] A. Deutsch, Mount Wilson Observatory, personal communication.

[53] M. Schwarzschild, Ann. astrophys. 12, 148 (1949).

knowledge the detailed mechanism of such oscillations cannot as yet be specified.

FIELDS IN RAREFIED COSMIC GASES

In dealing with this subject the writer must drop any pretense at even the most superficially complete coverage; he can at best discuss a few selected topics. Ionized cosmic gases range all the way from the ionosphere through the solar corona to interstellar matter in the galaxy. So far as the explanation of electromagnetic phenomena in these gases is concerned, little has as yet been achieved in the way of a thorough physical understanding. It will be our aim in the following to define the applicability of hydromagnetic theory as a valid approximation to a more elaborate theory of electromagnetic fields in ionized media (a general theory of plasma) which theory would be so complex as to be unmanageable. There is reason to believe that the hydromagnetic equations do constitute an excellent approximation in many important cases. In order to avoid confusion we shall in a few words repeat the significant features of this approximation as given in Part I.

Virtually all cosmic fluids are excellent conductors (the lower atmosphere of the earth and of some planets being a conspicuous exemption). Two facts of elementary electromagnetic theory are implied by this, namely, first that the displacement current is negligible, secondly that any accumulation of static charge disappears from the inside of a conductor with extreme rapidity (time ϵ/σ in mks units). The electric field can then be expressed in terms of the magnetic field, by

$$\mathbf{E} = -\mathbf{v}\times\mathbf{B} + \nu_m \nabla^2 \mathbf{B}. \qquad (1)$$

The field (1) is not in general divergence-free, but the field and the corresponding charges change quasistatically with the fluid motion provided the last term of Eq. (1) is small (large linear dimensions). The electrical potentials generated are, by Eq. (1), of order

$$\phi = BVL, \qquad (2)$$

where V and L designate a representative velocity and length, respectively. Finally, when Eq. (1) is valid the stress tensor reduces to its magnetic part, the electrical part becomes negligible.

Inductance is proportional to the square of the linear dimensions and hence becomes very large for bodies of cosmic size, leading to exceedingly large free-decay times for the magnetic fields. If an emf acts on a conductor it takes the same order of time to establish a steady-state current. Hence, for the same reason that inductive amplification becomes of paramount importance, the effects of any emf or electrical "polarization" of the medium become negligible in sufficiently large dimensions.

Let us now consider the relationship of hydromagnetic phenomena to the processes in the universe which are perceived in the form of radio noise. Sometimes magnetic fields might provide energy for radio noise, but it does not follow that a mathematical treatment of hydromagnetic fields must deal with the noise-generating processes. There is excellent evidence that a large part of solar and cosmic radio noise, especially the quiet-day noise from the sun, is due to free-free transitions in the spectrum of ionized hydrogen. Another source of noise can be the rotation of dust grains having an electrical moment.[54] Perhaps in rare cases such as the atmosphere of Jupiter which might be suspected to be an insulator, thunderstorms might be the cause. On the other hand, certain investigators have claimed that plasma oscillations generate noise of solar disturbances as well as cosmic noise. We need not enter into this problem at all, but merely remark that calculations show[55] that such plasma oscillations must needs occur in linear dimensions of the order of centimeters. Free-free transitions require very high temperatures, and plasma oscillations, if they are present, require electromagnetic excitation. But these processes must, as compared to hydromagnetic phenomena, appear only as irreversible sinks of energy; if they influence a hydromagnetic field otherwise than by absorbing its energy, it would only be through such thermal effects as the expansion of a heated compressible gas. There is no reason for presuming a close coupling of any other type.

In highly rarefied gases the electrical conductivity is modified by the presence of a magnetic field, owing to the coiling up of the carriers by

---

[54] The author is indebted to Mr. W. Erickson (University of Minnesota) for informing him of his analysis of this mechanism (to be published).

[55] A. Unsöld, Phys. Rev. **82**, 857 (1951).

the field. Cowling[56] has given the theory of this effect which consists of a general reduction in conductivity together with a "Hall"-current normal to both the impressed electric field and the magnetic field, making the conduction process anisotropic. The effects become large when the time between collisions is large compared to the period of gyration of an electron in the magnetic field. Schlüter[57] has given an alternative derivation of the effect which shows that one can use an isotropic conductivity throughout, provided one introduces an impressed "polarization" of suitable magnitude and direction.

We shall next have a look at magnetic fields in the solar envelope, since effects in the earth's outer atmosphere are clearly controlled from the sun. The solar photosphere where the visible spectrum originates is quite thin. If we include the reversing layer (where the absorption lines are produced) it is somewhat over 500 km thick. The chromosphere above it has a height of the order of 10 000–20 000 km. Beyond the chromosphere the corona extends out to not very well defined distances, often visible from one to several solar radii from the sun's disk. The layers named are only approximately defined and pass gradually into each other. Whereas in the photosphere the temperature decreases with increasing height, this trend is reversed in the chromosphere and the temperature begins to rise, eventually going much higher than the blackbody temperature of the sun's surface. This rapid rise continues as one goes into the corona. Over most of its thickness, out to at least twice the solar radius, the corona has a fairly constant temperature of approximately one million degrees.[41,58] It follows from simple principles of gas theory that layers where the temperature stays constant or even rises with height are extremely stably stratified. Turbulence is effectively suppressed and hence this ordinary vehicle of heat transport does not exist in the upper solar atmosphere.

The density is of the order of 10^{-8} cgs in the photosphere, or about 10^{15}–10^{16} P_c (where P_c designates protons, not necessarily ionized, per cm^3). It sinks to about 10^8–10^9 P_c at the bottom of the corona, and to about 10^6 at twice the solar radius. (To estimate the effect of magnetic fields in these gases it is convenient to note that the magnetic pressure is $B^2/8\pi$, where B is in gauss; the gas pressure is kTP_c.) We do not know what the magnetic field strength is in the upper atmosphere of the sun. Owing to the excellent electrical conductivity of the material and the exceedingly stable stratification of the region where the sharp rise in temperature occurs, it is rather senseless to extrapolate from the observations in the photosphere below.

The cause of the high temperature of the corona is unknown. There are numerous theoretical attempts[41] of which two will be mentioned. One of them assumes that the energy transport is by acoustical waves issuing from the convection zone (these would be supersonic shock waves in the upper atmosphere). Another hypothesis assumes that the waves traveling upwards are hydromagnetic.[42] In either case, a wave penetrating into a medium that thins out progressively will be strongly dissipated (after the manner of breakers at the ocean shore). It should be remembered that in any event the total mass of the corona is minute (say 10^{-5}–10^{-6} g/cm^2 of solar surface).

It is difficult to see how magnetic fields can penetrate the extremely stable strata of, or near, the chromosphere. If the layers were completely quiescent, penetration by virtue of molecular "diffusion" (that is, based on actual, rather than eddy, electric conductivity) would take thousands of years. It is not altogether impossible that this type of penetration takes place in the polar regions. There is here a clear contradiction between the hydrodynamic stability of a layer in which the temperature increases with increasing height and the observed violent dynamic processes at the same heights (e.g., spicules). There seems to be hardly any escape but the assumption that in these cases one is dealing with mass motions which originate in the top layers of the convection zone, and represent a bodily penetration (jet) into or across the stable layer. Such upward-moving masses would then carry their magnetic fields with them into the outer solar envelope. Unfortunately, nothing quantitative is known about magnetic fields in these outer layers, but there are observational indications for the existence of fields.

---

[56] T. G. Cowling, Monthly Notices Roy. Astron. Soc. **93**, 90 (1932).

[57] A. Schlüter, Z. Naturforsch. **5a**, 72 (1950); **6a**, 73 (1951).

[58] C. W. Allen Repts. Progr. in Phys. **17**, 135 (1954).

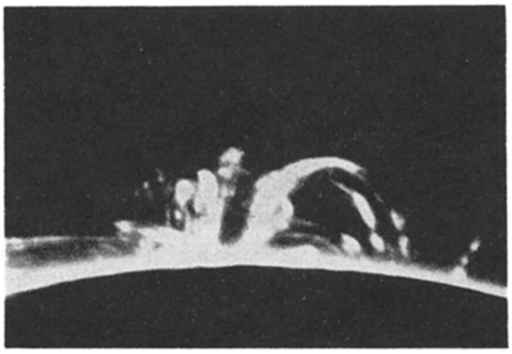


FIG. 13. Solar prominence. (Courtesy Dr. W. O. Roberts, High-Altitude Observatory, Boulder, Colorado.)

A streaky aspect of the corona (the streaks are usually designated as "streamers") is very pronounced around the poles of the sun. These "streamers" diverge from the polar caps, bending away from the solar axis in a manner which resembles closely the lines of force of a magnetic dipole. When this was first observed in the 19th century it gave rise to the presumption that the sun is a magnet. According to our present notions, however, the fields might be weak; a fraction of a gauss would probably suffice to explain the observed phenomena. The shape of the corona changes in time. Unfortunately, its detailed variation and its dependence on the sunspot cycle are still little known, since the outer corona is rather faint and streamers at large distances from the solar disk can only be photographed successfully at total eclipses. This restriction does not apply to the lower corona and chromosphere which are made visible by screening off the sun proper by a black disk after the method of Lyot.

The prominences (sometimes called protuberances) are often of great physical size (Fig. 13).[59] In front of the luminous solar disk they appear as dark "filaments." They are frequently associated with sunspots. Since they often show a pronounced arc structure they have been interpreted as "guided" by magnetic lines of force. They often change their details in a matter of minutes and hence the associated linear velocities are very large. Whether this represents an actual mass transport or merely some sort of wave motion traveling along the lines of force and changing the state of optical excitation, is not decided. Again, relatively stable prominences have been observed in association with photospheric magnetic fields. Some very impressive photographs of prominences and coronal streamers have been made, a number of which are now readily accessible.[41]

We turn now from the sun to the earth. The ionosphere has a variable conductivity depending on the variation of the supply of ultraviolet radiation from the sun. This phenomenon is of great practical importance for radio communications, but somewhat marginal in our context since it does not involve motions of the conducting matter. It is true that ionospheric winds have been found, and studied, during recent years, but these problems would lead us too far afield. The diurnal motions of the ionosphere on the other hand are conveniently measureable in terms of the diurnal component of the geomagnetic field. The interpretation is due to Chapman.[60] There are two main effects: the diurnal tide of the upper atmosphere and some convective motions occurring in the sunlit part of the ionosphere. Unfortunately the diurnal magnetic field at the earth's surface is compatible with an infinity of distributions with height, and hence the interpretation in terms of the phenomena quoted yields relatively little information about structure or dynamics of the ionosphere.

By far the most interesting terrestrial phenomena from the viewpoint of hydromagnetic theory are presented by magnetic storms and the closely associated auroras. The observational aspects of magnetic storms are briefly as follows. On magnetically "quiet" days the record of a magnetic observatory is quite smooth, with "noise" less perhaps than 1γ. (It is to be noted that the conventional magnetic recording instruments have damping times of the order of several minutes or more so that only relatively long periods are observed). On magnetically disturbed days the record is distinctly irregular and might show fluctuations of, say 10–40γ over a few hours. Among the disturbed days are found those where there is a major disturbance due to magnetic storms characterized by a fairly definite pattern. A magnetic storm is a world-wide phe-

---

[59] Courtesy, Dr. Walter O. Roberts, Director, High Altitude Observatory, Boulder, Colorado.

[60] S. Chapman and J. Bartels, *Geomagnetism* (Clarendon Press, Oxford, England, 1940), 2 volumes.

nomenon; its sudden commencement and its later features are simultaneous at observatories all over the world, the approximate average behavior being about as follows: The horizontal intensity first rises quite rapidly by, say, 20γ and then falls somewhat less rapidly. After about 6–8 hours the anomaly has become negative and remains so for the remainder of the storm. This negative variation reaches its extremum some 12–16 hours after commencement, its value ranging from say 30–50γ for weaker storms to several hundred γ for exceptionally severe storms. Thereafter the deviation from the normal field decays very gradually, over several days. The vertical component of the field shows somewhat similar though much less pronounced variations.

"Of the many theories of magnetic storms advanced during the past fifty years, there survives only that of Chapman and Ferraro.[60,61] This theory postulates an electrically neutral, ionized stream of particles expelled from the sun with velocities of the order of 1000 km/sec, in accordance with the well-known fact that particles emitted during solar catastrophes appear to take about a day to reach the earth's neighborhood. Chapman and Ferraro find that the surface of this stream, which is a good conductor, will be retarded by the earth's magnetic field at a distance of several earth's radii, and that a region around the earth will be 'forbidden' to the solar stream. In this space the earth's field will be slightly compressed, thus producing the increase in field strength which is generally characteristic of the first phase of magnetic storms of sudden commencement. The evidence is strong that after some hours a westward-flowing electrical current is set up in a ring lying in and near the (geomagnetic) equatorial plane, at a distance of several earth's radii. Chapman and Ferraro were unable to account satisfactorily for the formation of this ring current, but made a careful examination of the (quasi-) steady state of this ring, once in existence. . . . The strength of their theory lies in the extreme care with which each necessarily occurring process has been analyzed, often by several different approaches."

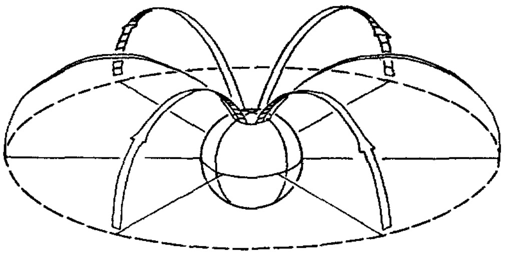


FIG. 14. Paths of auroral primaries from magnetic equatorial plane.

The preceding passage, taken from an article by Martyn[62] expresses better than the writer could do the importance of what is undoubtedly the earliest systematic theory in the subject with which this review is concerned. The magnetic dipole field is weakest in its equatorial plane, and this is where we would expect the closest penetration of the stream into the "hollow" carved out by the earth's field. In a very crude approximation we would expect conducting matter to flow along the magnetic lines of force of an external, sufficiently strong field, since this fails to set up an induced current and hence there are no reactive mechanical forces. In a next approximation $\mathbf{v}$ and $\mathbf{B}$ are no longer parallel; using the component of $\mathbf{v}$ in the meridional planes we then obtain an electric current in the ϕ direction by (3, Part I). Martyn[62] has given a simple quantitative estimate of the distance of the ring current, he finds its mean radius to be about 5.5 earth's radii. This value is rather insensitive to changes in density of the ionized matter. Martyn finds $P_c \sim 20$ for the density, which is somewhat larger than Chapman's earlier value of $P_c = 1$ but might still be an underestimate.

The auroras,[63,64] are closely correlated in their incidence with magnetic storms and there is no doubt of an intimate physical connection. The auroras are concentrated in a ring centered about the magnetic pole, the line of maximum occurrence is nearly a circle of opening angle ~23° about the magnetic pole. This applies to the northern hemisphere; in the southern hemisphere the phenomena are similar qualitatively, though with some quantitative variations. It is generally admitted that the primaries causing auroras

---

[61] S. Chapman and V. Ferraro, Terrestrial Magnetism Atm. Elec. **36**, 77, 171 (1932); **37**, 147, 421 (1933); **45**, 245 (1940).

[62] D. F. Martyn, Nature **167**, 92 (1951).

[63] V. C. A. Ferraro, Phil. Mag. **2**, 265 (1955).

[64] See the article by Chamberlain and Meinel in *The Earth as a Planet*, G. Kuiper editor (Chicago University Press, Chicago, 1954).

travel toward the earth nearly along magnetic lines of force (Fig. 14). As Martyn[62] points out, the lines of force going through the equatorial planes at 5.5 earth's radii meet the earth's surface at an angle of 25° from the magnetic pole. This is a sufficiently close agreement to give more support to the idea that the primaries of auroras are part of the same ionized material streams which produce the magnetic storms.

A direct spectroscopic determination of the speed of the primaries was carried out by Meinel[65] who observed the Doppler shift of the Balmer line H_α in auroras. Looking perpendicularly to the magnetic field there is no shift but only a broadening. Looking along the field one finds a strong displacement, the mean shift corresponding to a velocity of arrival of 450 km/sec. The line extends asymmetrically to much higher velocities; at 3000 km/sec it still has 10% of its peak value, and Meinel assumes that the velocity of the primaries before hitting the earth was close to 4000 km/sec, attributing the lower velocities to slowing down in the atmosphere.

We are induced to think that these gaseous streams penetrating into the upper atmosphere must have undergone two accelerations: first, one at the sun; the correlations of geomagnetic phenomena with preceding violent events at the sun indicate speeds of travel of 500–1500 km/sec. This figure is of course vastly in excess of the average velocities of mass motions observed on the sun. There must then be a further acceleration of these gases near the earth to account for the high velocity found by Meinel. It is not to be presumed that these accelerations are caused by electrical fields; they should be interpreted as hydrodynamical (hydromagnetic) in nature, to be compared perhaps to jet formation in ordinary hydrodynamics. We know for example that quite weak temperature gradients in the upper atmosphere of the earth become concentrated and lead to the acceleration of a small fraction of the air so as to produce the remarkable phenomenon of the stratospheric jet stream whose speed is a multiple of any other wind speed. There seems to be no reason, so far as the author can see, to doubt seriously that the velocities related to magnetic storms and auroras are true fluid velocities whose origin can be described by purely "classical" dynamics of compressible, conductive gases. The theory will still have to be developed.

The explanation of auroras has long labored under another, somewhat similar difficulty. The visual displays of auroras reach down to a height of about 100 km above ground, sometimes they come close to 80 km. This means that the aurora has penetrated, counting from the top of the atmosphere on down, the equivalent of from one to several cm of air at standard pressure. To penetrate 1 cm of standard air, a proton needs an energy of $7\cdot10^5$ ev, over ten times the energy deduced from the spectroscopic observations. It seems unwise to conclude that the primaries of the auroras have energies of this extremely high order; it is much more likely that the spectroscopic observations give the true speed of the incident ionized matter. In this case the deep penetration into the atmosphere would be a secondary phenomenon, perhaps some sort of compression generated by the impact of the incident stream.

The problem of ionized gas and related magnetic fields in the solar system at large is difficult since it requires data on rather tenuously distributed matter. It has long been known from studies of the Zodiacal light that there are dust particles in the solar system. Recently Behr and Siedentopf[66] have succeeded in separating quantitatively the unpolarized component of the scattered light, attributed to dust, from the polarized component attributed to free electrons. They indicate a density of about 10^3 electrons/cm^3 at a distance from the sun of 0.6 a (where a is the radius of the earth's orbit) falling to 6×10^2 at distance a and then rapidly decreasing to 1.2×10^2 at 1.3 a. An interesting partial corroboration of these results comes from the work of Storey[67] on the propagation of atmospheric "whistlers" which appears to take place along the lines of force of the geomagnetic field. An estimate of the electron density at a height of two earth radii leads to a value of $4\cdot10^2$/cm^3. It is only to be hoped that such valuable observations as the ones quoted above can soon be verified further.

If the ion densities quoted for the solar system are real, they could clearly carry a rather sub-

---

[65] A. B. Meinel, Astrophys. J. **111**, 155 (1950).

[66] A. Behr and H. Siedentopf, Z. Astrophys. **32**, 19 (1953).

[67] L. R. P. Storey, Trans. Roy. Soc. (London) **246**, 113 (1954).

stantial uniform magnetic field provided they would remain undisturbed for some time. The indications are that the latter is not the case. The observations on magnetic storms and auroras do not seem to admit of any simple explanation other than that of streams of ionized gas ejected from the sun which travel at speeds of 500–1500 km/sec, corresponding to transit times from sun to earth of $\frac{1}{2}$–$\frac{3}{2}$ days. Just how these streams manage to reach the earth in the presence of the stationary gas density indicated is not quite clear and has apparently not been investigated; one must probably think of jets threading their way through a tenuous medium rather than of gas clouds expanding into empty space. Biermann[68] found that the well known bending away from the sun of the comets' tails (formerly erroneously attributed to radiation pressure) is in all likelihood the result of pressure exerted by gas ejected from the sun. Quite recently Davis[69] has pointed out an important astrophysical consequence of the ejection of streams of matter from the sun, related to the over-all galactic field (see the following). The net average effect of such ejection would clearly be a pressure exerted upon the gas filling the galaxy at large. As a consequence the magnetic field lines of the galactic field will be pushed apart, forming a sort of "hollow" with the sun at its center. The magnitude of the hollow may be estimated as of the order of 10^2a–10^3a. When the gas coming from the sun collides with the "walls" of the galactic hollow it will be stopped. Since such gas undoubtedly carries its own magnetic field we might expect some fairly complicated hydromagnetic phenomena in the outer parts of the solar system at distances fairly large compared to the size of the earth's orbit. The galactic hollow is so far hypothetical, but it is likely to be detected sooner or later by its influence upon cosmic rays.

We now come to the galactic magnetic field. In 1949 Hiltner and Hall discovered[70] independently that starlight which has undergone absorption in its passage through the galaxy becomes partially polarized. Two theories of this phenomenon have been given, both of them agreeing that the polarization is due to the alignment by a magnetic field of particles in galactic dust clouds. One theory[71] assumes ferromagnetic particles; the other view[72] attributes the alignment to a paramagnetic relaxation effect in grains which are spinning very rapidly. We cannot discuss these interesting theories here, especially since they agree in their main thesis, namely that the optical polarization effect is ultimately due to a galactic magnetic field. The earth is located in a spiral arm of the galaxy. The magnetic field runs longitudinally in the spiral arm, with some local variations in direction. Clearly, the above-mentioned theories of the orientation of grains do not readily yield quantitative values for the field; they do, however, go far enough to show that the field strength deduced from other arguments is compatible with the optical observations.

The density of the galactic gas is generally estimated as of order $1P_c$. This makes the total amount of gas in the galaxy less than, but comparable in order of magnitude with, the matter condensed in stars. This gas seems too extensive to have been ejected from stars; probably it has never been condensed in stars. If one assumes equipartition between the kinetic energy of the gas and the magnetic field (Part I) a value near 10^{-5} gauss is obtained.[71] Again, according to Chandrasekhar and Fermi[73] the magnetic field in the galaxy keeps the gas from collapsing by gravitation, and from this argument a field strength of 6×10^{-6} gauss is derived. The two estimates are of course related, and unless the dynamical interpretation given to the field is grossly in error, the galactic field must be of this general order.

The existence of a magnetic field in the galactic medium has not yet been explained. Since the galaxy is a rapidly rotating system (its kinetic energy of rotation being very large compared to the energy of the random motion of stars) an analogy with the toroidal field in the earth and sun suggests itself at once. It is not easy, however, to apply the two-stage feedback process described in Part I. This process requires motions across the lines of force of the toroidal field. Now there is no convection in a direction normal to

---

[68] L. Biermann, Z. Astrophys. 29, 274 (1951).

[69] L. Davis, Jr., Phys. Rev. (to be published).

[70] See references 66 and 67 for detailed literature.

[71] L. Spitzer and J. Tukey, Astrophys. J. 114, 187 (1951).

[72] L. Davis, Jr. and J. Greenstein, Astrophys. J. 114, 206 (1951).

[73] S. Chandrasekhar and E. Fermi, Astrophys. J. 118, 113, 116 (1953).

the extension of a spiral arm. In a cylindrical system held together by gravity the component bodies will tend to oscillate about the cylinder axis. Whether this type of motion, usually assumed for stars, also occurs for gas clouds and whether it can play a role in maintaining a galactic magnetic field, remains to be investigated.

In the preceding pages we have tried to give a survey of various phenomena in the earth and universe which can be described in terms of hydromagnetic theory. In concluding we note two omissions. The first concerns the application of hydromagnetic theory to arguments concerning the origin of the solar system, a line of study pursued by Alfvén.[74] Since cosmogonic theories are still in a highly preliminary stage, it will suffice to refer the reader here to Alfvén's book. The second omission is that of theories which try to explain the origin of cosmic rays on the basis of cosmic hydromagnetic fields. To give an exposé of this problem would double the length of the present article. We shall only point to the pertinent observations and theoretical arguments which have been gathered by Biermann[75,76] and his associates (see also Teller).[77] If hydromagnetic fields are responsible for the acceleration of very fast particles, the order of magnitude of the voltages to be expected is given by Eq. (2) above. Really high potentials corresponding to the upper part of the cosmic-ray spectrum can then only be achieved in extremely large dimensions, of galactic order. The well-known mechanism of Fermi[77] accelerates particles by collisions with moving inhomogeneities in the galactic magnetic field. On the other hand it seems that frictional dissipation of such irregular fields is intense[78] so that only a fraction of the field energy can be used to accelerate cosmic rays. The energetic difficulties then become most serious; the energy consumed by the galactic gas is so large that it could, for instance, upset the gravitational equilibrium of a spiral arm in a time relatively short on the cosmic scale. On the other hand, the galactic "hollow" of Davis[69] could retain cosmic-ray particles of the lower energies in the solar system, long enough to impart appreciable energies to them by acceleration in hydromagnetic fields. But the faster cosmic-ray particles cannot be so retained by any magnetic field that may reasonably be assumed. Thus the now fairly old controversy of solar-system *versus* galactic origin of the cosmic rays still remains unresolved. There is no basic reason, of course, for assuming that the cosmic rays owe their origin to one single machanism of acceleration rather than to several. But up to now no discontinuity in the cosmic-ray spectrum has been observed such as one would suspect to exist in the case of multiple accelerating mechanisms.

---

[74] H. Alfvén, *On the Origin of the Solar System* (Clarendon Press, Oxford, England, 1954).

[75] See Chapter I, edited by L. Biermann, in the book *Kosmische Strahlung*, W. Heisenberg, editor (Springer Verlag,Berlin, Germany, 1953).

[76] L. Biermann, Ann. Rev. Nuclear Sci. **2**, 335 (1953).

[77] E. Teller, Repts. Progr. in Phys. **17**, 154 (1954).

[78] E. N. Parker, Phys. Rev. **99**, 241 (1955).

The most fundamental aim of education is perhaps the development of the power of good judgment in answering the questions which cannot be answered by science, but which must be answered in life—problems which no intensification of scientific training by itself will solve. Unless the student acquires not only scientific understanding and skill in scientific method, but also learns to practice and develop the powers of nonscientific judgment and decision, the university has failed in its task of producing an educated man or woman.—Editorial, *Nature*, October 8, 1955.