17 September 1965, Volume 149, Number 3690 SCIENCE

Solar Magnetic Fields

Study of solar magnetic field shows the importance of large- and small-scale structure in solar activity.

V. Bumba and Robert Howard

The intriguing problem of solar activity and its many manifestations has been considered for several decades to be closely associated with the existence of magnetic fields on the sun's surface. Hale (*1*), at Mount Wilson, was the first to demonstrate the presence of magnetic fields on the sun. These were the strong magnetic fields of sunspots. He and his collaborators later spent many years in an attempt to find a weak general magnetic field on the sun. Because of the insensitivity of the photographic technique—the only way open to them at that time—their results are considered inconclusive. Hale (*2*) himself was associated with the first attempt to make a photoelectric measurement of the sun's magnetic field. In 1952 H. W. Babcock (*3*) developed the solar magnetograph, a sensitive photoelectric instrument for measuring solar magnetic fields. Solar magnetographs are in use now in several observatories around the world in essentially the form first developed by Babcock.

Measurement of Solar Magnetic Fields

Practically all measurements of magnetic fields in the solar atmosphere utilize the Zeeman effect. Magnetic fields in sunspots are measured in hundreds of gauss, and visual or photographic techniques with appropriate analyzing optics give results accurate to within 100 gauss or so. Magnetic fields outside sunspots are, in general, weaker by an order of magnitude. With good angular resolution, photographic techniques can detect the stronger fields in plages or in the enhanced network, but photoelectric techniques are required for studying the weaker fields (<50 gauss).

Except in the case of some recent work at two observatories in the Soviet Union (*4*), all magnetographs measure only the longitudinal component of the solar field. Figure 1 is a drawing of the absorption line and the magnetograph slits. By means of an electro-optic device the light entering the spectrograph is modulated in such a way that light of alternately one and the other circular polarization is admitted to the spectograph. The effect at the magnetograph exit slits is that of alternately admitting one and the other longitudinal Zeeman component. Thus the slits which are on opposite wings of the line "see" an alternating signal. The signal from each slit is 180 degrees out of phase (of the modulation signal) with the signal from the other slit. The difference between the two signals (as determined in a difference amplifier) is proportional to the strength of the magnetic field as long as the splitting is not so great as to move the line to a point where the shape of the line profile gives a response which is no longer a linear function of the separation of the Zeeman components. Because of the subtraction, this method of measuring the magnetic field is insensitive to the effects of instrumental polarization. In the case of the transverse fields, the Zeeman components that are split by the field have the same polarization, and the subtraction method will not work; thus, in the measurement of the transverse component, the instrumental polarization (which in most instruments is considerable) must be measured and somehow eliminated. For this and other reasons the sensitivity of magnetographs measuring the transverse component is lower by about two orders of magnitude than that of magnetographs measuring the longitudinal component. The actual sensitivity of magnetographs, of course, depends upon the size of the entrance slit, the integration time, and other factors, but in general a sensitivity of about ½ gauss is attainable in longitudinal scanning of an active region.

A technique developed by Leighton (*5*) uses the principle of the magnetograph to obtain photographs of the magnetic fields on the sun. A beam splitter is used to give two images of the sun in oppositely polarized light on the slit of a spectroheliograph. The two spectroheliograms that result differ in that a magnetic field of a given polarity appears darker than the background on one image and lighter than the background on the other. A photographic subtraction process is used to make the final picture. In general this photographic process is somewhat less sensitive than the photoelectric technique of the magnetograph, but it has the great advantage of presenting a two-dimensional picture of the magnetic field configuration in a relatively short time. With a magnetograph such a picture can only be built up by means of a raster scan of some sort.

Dr. Bumba is on the staff of the Astronomical Institute of the Czechoslovak Academy of Sciences, Observatory Ondrejov, Ondrejov, Czechoslovakia. Dr. Howard is on the staff of the Mount Wilson and Palomar Observatories, Pasadena, California.

Sunspots

All sunspots contain strong magnetic fields that are more or less radial to the sun. Systematic measurements of the magnetic fields of sunspots cover a period of nearly 50 years. One of the interesting properties of sunspots is that they always occur in groups. The distribution of magnetic polarities among the spots of a group follows a well-established pattern. Most groups at some time during their life consist of a single spot, and some groups have an irregular distribution of spot polarities and are classified as magnetically complex, but in general a sunspot group is divided into two halves on an axis which lies roughly east-west on the sun. The spots in the preceding part of the group (in the sense of rotation—that is, west) have magnetic polarities which are the same, and opposite to those of the following spots. In groups in the northern hemisphere of the sun the preceding spots are all of the same polarity during one 11-year activity cycle, and this is the opposite of the polarity of the preceding spots in the southern hemisphere. During the next 11-year cycle these polarities are all reversed. Thus, magnetically, the solar activity cycle is a 22-year cycle.

Sunspots are cooler by many hundreds of degrees Celsius than the surrounding photosphere. The fact that they are not as bright, when viewed in white light, as the photosphere is evidence of this fact. There is other evidence, such as the relative weakness of absorption lines of ionized elements in sunspot spectra. Earlier theories (*6*) of sunspots usually attempted to explain the presence of magnetic fields in the cool spot by some means, such as generation of the fields by vortex motion of ionized particles around a region of low pressure. In more recent times more has been learned about the behavior of plasmas in a magnetic field, and recent theories have dealt with explaining the low temperature as a consequence of the presence of strong magnetic fields. It is generally believed now that convective motions, which account for almost all the outward flow of radiant energy for some distance below the visible surface of the sun, are effectively impeded in the strong vertical magnetic field of a sunspot (or at least that the modes of motion which are most efficient in the transport of energy are impeded). With no energy received from below by convection, the temperature of the sunspot decreases until the energy lost by radiation from its surface is in equilibrium with the energy gained by radiation transfer from below and from the side, or by other means.

This is not intended to imply that sunspots are thoroughly understood. A complete theoretical understanding of the sunspot phenomenon may not come for some time. One difficulty is that the observational problem of getting accurate photographs, photometric data, or spectra is enormous because of the ubiquitous scattered light from the earth's atmosphere and from the telescope (*7*).

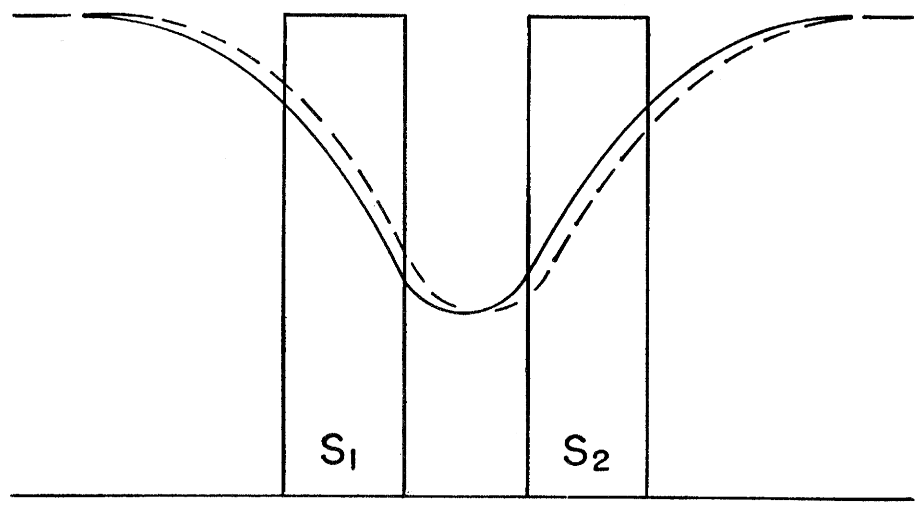

Fig. 1. Schematic orientation of magnetograph exit slits with respect to the solar absorption line profile. The solid and dashed profiles represent the two components of the simple longitudinal pattern. The vertical lines show the position of the exit slits.

Active Regions

A sunspot group is really just the manifestation, in white light, of a larger phenomenon. To a very great extent in this field, as in all fields of astronomy, we are limited by our observations, and our outlook on a problem is shaped by the particular ways in which we happen to be able to observe the phenomena. These ways of looking at the phenomena often originated in the past, for reasons of necessity, and are not now the most logical ways. A good example is the phenomenon of the solar flare. Flares have been observed almost exclusively in the Hα line, and the phenomenon has often been defined simply as a sudden brightening in that line. More recently it has become clear that the flare observed in Hα is one manifestation of a much larger event involving the release of energy in many parts of the spectrum, including radio frequencies and x-rays, and that much of this energy is released at locations other than that of the flare observed in Hα. Nevertheless, we still frequently discuss flares as if all that was involved was a brightening in Hα. Observation of active regions has a somewhat similar history. For centuries all evidence of solar activity rested on white-light counts of sunspots and, more recently, measurements of faculae. In the last 75 years or so it has been possible to obtain spectroheliograms, and thus to get some idea of the nature of an active region in the chromosphere. In the last decade it has been possible to map magnetic fields on the sun's surface and thereby add a really basic physical quantity to our picture of the solar atmosphere. Nevertheless, we still talk about sunspots as if they *were* the active regions.

An active region usually, although not always, has a sunspot group associated with it during some part of its lifetime. The basic visible feature of an active region is the plage—the emission seen best in the Hα line or the H or K line of calcium II. Figure 2 shows spectroheliograms taken in light of the Hα line and the K line of calcium II and a photograph taken in white light. It has been established that more or less radial magnetic fields of the order of 100 gauss occupy the position of plages. Thus it is probable that, as with sunspots, the basic feature is the magnetic field and the visible phenomena follow from the presence of the field. It is generally believed that the higher temperature evident in the plage is the result of heating in the chromosphere by dissipation of magnetohydrodynamic waves, which occur because of the presence of the magnetic field (*8*).

Active regions form relatively rapidly, usually growing to full size within 2 to 5 days of the first appearance of the plage. The actual growth of an active region begins at a concentration of magnetic flux associated with the calcium network, in an area of scattered magnetic fields of an old active region (*9*). In general, the plage forms first at one end of the final region (usually at the "following" end), and expands to the other end. It sometimes happens that during the first day the newborn plage is observed to have only one magnetic polarity. The sunspot group appears within a few days of the birth of the region. The newborn plage is relatively bright, with a smooth boundary. As the plage ages, the boundary appears more ragged, and parts of the emission appear to break off. The area surrounding the plage soon gets filled up with these little bits of emitting material. Eventually this process eats away all the plage, and all that remains is the collection of bright specks.

Leighton (*10*) has suggested that the breakup of active regions is due to the effects of the supergranular motions—the motions in large cell-like structures (30,000 km in diameter) which occupy all the quiet surface of the sun. Horizontal motions, outward from the center of the cells, are observed to be of the order of 0.5 km/sec (*11*). In the light of the K line of calcium the cells are roughly outlined by small emission patches. This is the calcium network referred to above. These emission patches are observed to be concentrations of more or less radial magnetic fields (*12*). Lines of force are presumed to be swept outward by the supergranular motion and to be concentrated at the boundaries of the cells (see Fig. 3).

When observed in the light of the $H\alpha$ line the sun is seen to have other interesting features associated with active regions (see Fig. 2). Long dark features called "filaments" lie apparently peacefully in quiet parts of the sun or in active regions. These features are the same as the prominences seen off the limb of the sun (Fig. 4). It is observed that filaments lie almost always between regions of opposite magnetic polarity (there are a few apparent exceptions in complicated situations within active regions). It is

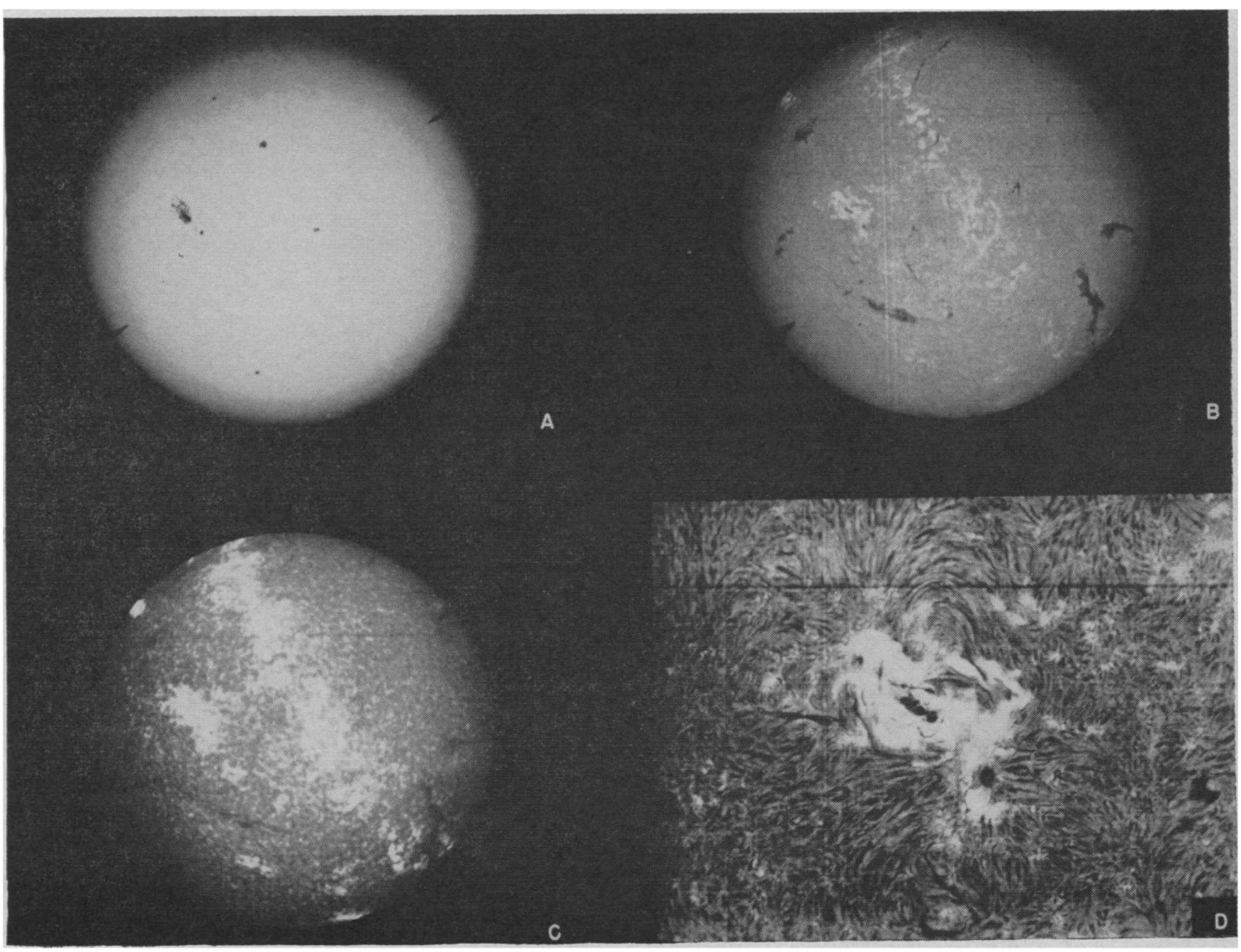

Fig. 2. Four views of the sun, 15 September 1949: (*A*) whole disk photographed in white light; (*B*) whole-disk spectroheliogram taken in light of the $H\alpha$ line; (*C*) whole-disk spectroheliogram taken in light of the K line of calcium II; and (*D*) enlarged section showing an active region, in light of the $H\alpha$ line.

clear that the magnetic field plays some role in the support of prominence material. One view is that the weight of the prominence material sags the arching lines of force, forming a trough within which charged particles of the coronal gas are trapped (*13*). In the neighborhood of the plage photographed in light of Hα there are concentrated many small elongated dark features, which can be seen in Fig. 2. These "fibrils" can also be seen scattered in quiet regions where they are not so neatly organized. Such features are associated with the spicules, which are small bright jets of matter seen shooting above the limb of the sun and falling back. Within an active region the fibrils are aligned parallel to the gradient of the magnetic field.

Our understanding of the chromosphere is by no means complete. The role of the magnetic field in the structure and thermodynamics of the chromosphere remains basically unsolved. In recent years interest in the appearance of the chromosphere on the disk, particularly as this relates to motions of the chromospheric material, has increased. Athay has recently reviewed the field of chromospheric physics (*14*).

The Decay of Active Regions

Soon after the development of the solar magnetograph, H. W. and H. D. Babcock (*15*) investigated the distribution of magnetic fields on the solar surface. They found that there was a more or less well-defined polar field, seen above about 60° latitude, positive in the north and negative in the south. At lower latitudes they could see bipolar magnetic regions, which corresponded with active regions, and unipolar magnetic regions, which did not correspond with any obvious chromospheric features. In general these unipolar regions were larger than the bipolar regions and were made up of weaker fields. At the time of the last activity maximum, H. D. Babcock (*16*) observed that the polar fields reversed their polarity—first the south pole, then, about 18 months later, the north pole. This led H. W. Babcock (*17*) to theorize that many of the characteristics of solar activity—such as the latitudinal drift of sunspot activity during the cycle, the preferential tilt of the axes of spot groups, and the variation of the frequency of occurrence of sunspots during the cycle—could be explained in the following way. If the cycle begins with a shallow, submerged, axially symmetric magnetic field, this field will become drawn out by the shearing effects of differential rotation and a spiral configuration will be produced. Ropes of flux lines will form, with local concentrations that bob to the surface to form bipolar magnetic regions. As the flux ropes are used up in this way, the preferential latitude for spots drifts equatorward because of the form of the differential rotation. Dying bipolar regions split apart; the following (easternmost) portions migrate to the poles and the preceding portions move toward the equator, where they cancel fields from the opposite hemisphere. Thus the stage is set for the next cycle, with polarity of the polar fields reversed.

Observations of solar magnetic fields have continued at Mount Wilson, and recently some attention has been given to magnetic data accumulated over several years (*18*). In these full-disk magnetograms an angular resolution of 23 seconds of arc was used, compared to 70 seconds of arc for the earlier observations studied by the Babcocks. In broad outline, the recent results confirm the earlier results; however, the increased resolution has made a considerable difference in the amount of detail that can be seen, and we are now presented with a complex picture of weak magnetic fields that cover a large portion of the sun's surface.

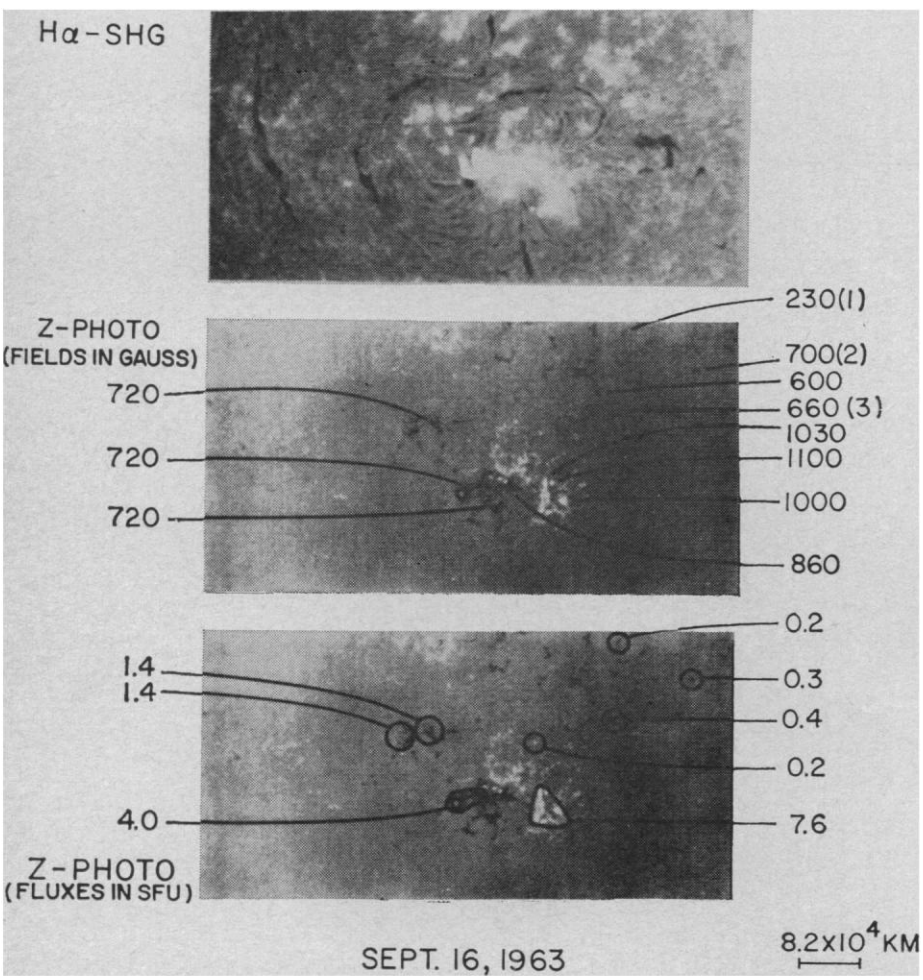

Fig. 3. (Top) Spectroheliogram taken in the light of the Hα line (Hα-SHG). (Middle and bottom) Two identical magnetic spectroheliograms (Z-photos) of the same region, one showing magnetic field strengths, the other showing magnetic fluxes in arbitrary units (solar flux units, or SFU). A tone lighter than the background indicates positive polarity; a tone darker than the background indicates negative polarity. Note that even small emission features in the spectroheliogram at top have strong magnetic fields associated with them. [Photograph courtesy of N. R. Sheeley]

Figure 5 is a synoptic chart of weak solar magnetic fields. A large portion—perhaps over half—of the sun's surface is covered with magnetic fields which, as measured with the 23-second-of-arc resolution of the magnetograph, are stronger than 2 gauss. This pattern shows remarkable regularity in the alternating pattern of polarities. The inclination of these large features is caused by the differential rotation of the sun—lower solar latitudes rotate with a higher angular velocity than higher latitudes do. We have called these weak fields the "background fields." The lifetime of the background-field pattern is not easy to estimate because the pattern is constantly undergoing slow changes. Gross features may be considered to have, in general, a lifetime of many months. On this background of weak fields the strong fields of active regions form, and then, after a few days or a few weeks, depending on the size of the region, the magnetic fields of the active regions begin to merge into the background field pattern. The pattern itself is made up of the magnetic fields of old active regions. Figure 6 is a synoptic chart from the period of minimum solar activity. The pattern of background fields is very weak where it is seen at all.

It is possible to see some degree of large-scale order in the appearance of active regions. During rather quiet periods of the activity cycle it is apparent that active regions occur over a period of some months in groups or "complexes" of activity. That is, a certain density of activity is maintained over a zone of 90° or so in longitude. In general, the first-appearing active regions of the complex tend to be larger than the last-appearing regions. The effect of the fact that the active regions are crowded into some longitude zone is to produce a pattern of background fields. During the years of maximum activity, the complexes of activity, if there are any, are so crowded together that a great deal of overlap must occur, and the result is that individual complexes cannot be distinguished. The larger complexes of activity have associated with them large regions of "following" polarity for that hemisphere, stretching eastward on the poleward side of the background fields. These regions are called unipolar magnetic regions. One cannot, perhaps, strictly speaking, prove a causal relationship between the active regions of the complex of activity, but, because of the striking appearance of the background field pattern and the frequent appearance of the unipolar magnetic region, it seems reasonable to speak of the complexes as if they are real.

Thus the unipolar magnetic region is a part of a very large pattern of magnetic fields of both polarities; the background fields seem to be of predominantly "leading" polarity for that hemisphere. During the period of the recent study most of the activity, and therefore most of the unipolar regions, was in the northern hemisphere. Consequently, in general, the net background field was predominantly of positive polarity. There are other large magnetic features associated with the same pattern of background fields which results in the formation of a unipolar magnetic region. Preceding the unipolar regions, and of opposite polarity, is a large feature which we have called the "ghost unipolar magnetic region." The magnetic fields of a ghost region are very weak and scattered and can be seen only on magnetograms of the finest quality. The field averaged over the region is less than 2 gauss. The magnetic flux from the ghost region is less by at least a factor of 5 than that of the unipolar regions. In the hemisphere opposite to that of the unipolar and ghost unipolar magnetic regions there are often very weak features that mimic in appearance the unipolar region and its ghost. Thus the large-scale pattern, which seemed to repeat itself nine times during the 5-year interval studied, was as follows: background fields at lower latitudes (the upper limit in latitude depending upon phase in the cycle); unipolar magnetic regions stretching poleward and eastward of the center of the background field pattern; ghost regions preceding the unipolar regions; and weak mirror images of the unipolar and ghost regions in the opposite hemisphere. The whole pattern of each complex stretched over a longitude range of perhaps 180°. Some features of the pattern may be seen in Fig. 5.

The unipolar magnetic regions moved poleward, as Babcock (*17*) predicted

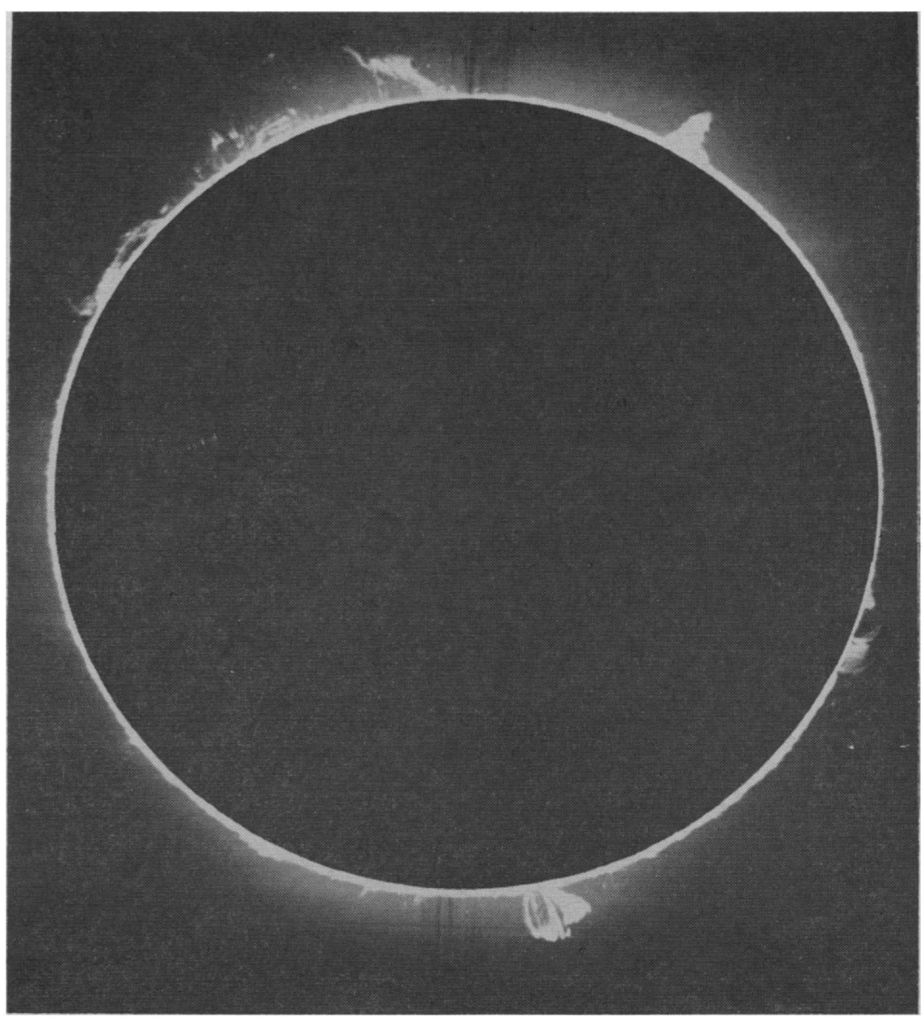

Fig. 4. Prominences shown at the limb in a spectroheliogram taken in light of the K line of calcium II, 9 December 1929.

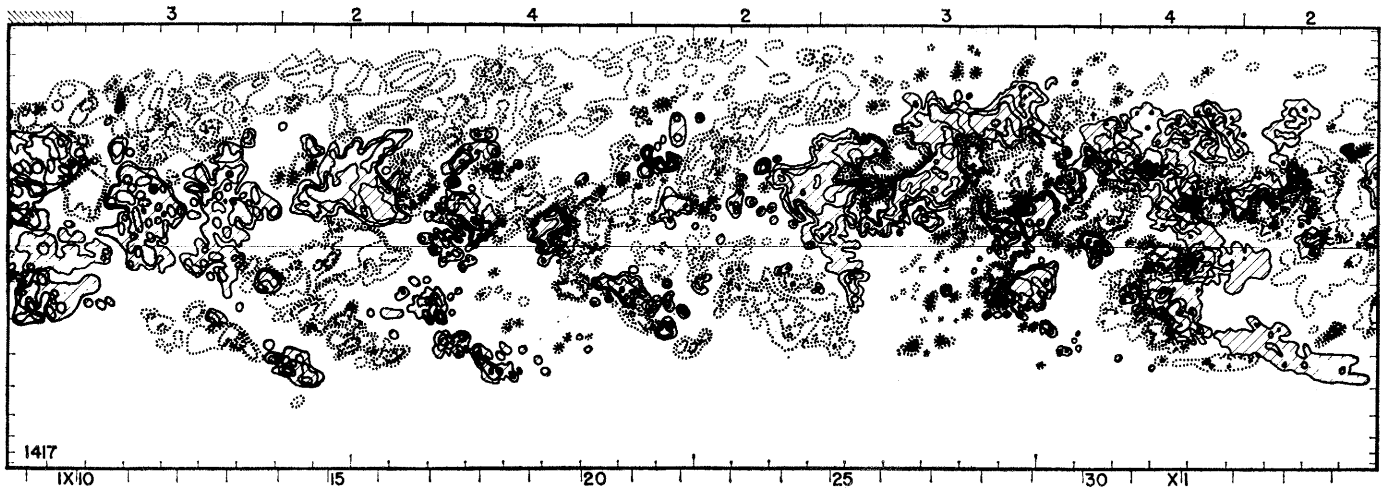

Fig. 5. Synoptic chart of solar magnetic fields for solar rotation number 1417 (August 1959). This is a rectangular equal-area projection. Solid lines and hatching represent positive polarity; dotted lines and shading represent negative polarity. The first isogauss level is 2 gauss. The equator is drawn across the center, and every 10° in latitude is marked at the sides. The scale at top gives an indication of the quality of the magnetograms from which the synoptic chart was drawn, with 4 the best; the hatching at left end of the scale represents an area drawn more than 40° from the central meridian of a magnetogram. The date (Roman numerals) and longitude are given on scales at bottom.

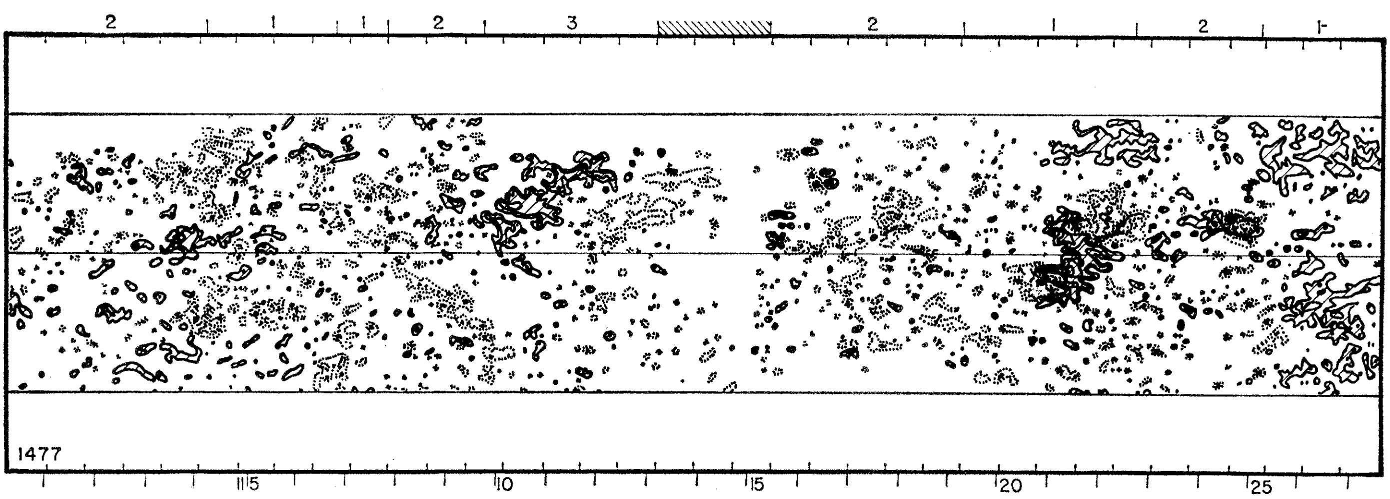

Fig. 6. Synoptic chart for solar rotation number 1477 (February 1964). Designations are the same as for Fig. 5. Observations beyond ±40° are not included.

they would. This is undoubtedly the mechanism responsible for the reversal of the polar field of the sun, and we may confidently predict that at the next activity maximum the polar fields will reverse again. There is a problem about the amount of flux which has been observed to drift toward the north pole of the sun since the observed reversal of the north polar field. Each unipolar region has a flux of about 10^{22} maxwells. This means that a flux of about 10^{23} maxwells has drifted toward the north pole since the reversal at maximum.

Even when generous allowance is made for geometrical foreshortening and for a large amount of flux of opposite magnetic polarity in ghost unipolar regions and in other features that may have drifted to the pole undetected, there still remains a discrepancy between the increase in the north polar field which would be predicted and the observations, which indicate, if anything, a decrease in the strength of the north polar field in the years since maximum. There is a similar sort of problem in the south, where the strength of the polar field has decreased enough to bring the field below the limit of detectability for the last several years, with no poleward drift of features of the opposite polarity apparent. Thus, unless we are being fooled by large-scale motions of magnetic features too weak to be detected, there is some mechanism at work which dissipates large-scale fields at a rate more rapid than theory can account for. The solution may lie in the fine-scale structure of solar magnetic fields. Some work has been done in the last few years on this observational problem, and as observational techniques improve we may look to this field for new advances in our understanding of solar magnetism.

Solar Magnetic Fields and Geomagnetic Activity

Since the development of the solar magnetograph 13 years ago, great strides have been made in our understanding of the distribution of magnetic fields on the surface of the sun, and, as a result of recent comparative studies of magnetographic, satellite, and geomagnetic data, in our understanding of the solar origin of magnetic fields near the orbit of the earth. The existence of a very-large-scale coherent pattern of solar magnetic fields, both across the surface of the sun and out into interplanetary space, is an extremely important finding. It promises to open new directions of research in the fields of solar activity, solar terrestrial relationships, and stellar magnetic fields.

References

1. G. E. Hale, *Astrophys. J.* **28**, 315 (1908).
2. "Ann. Rept. Mt. Wilson Obs.," in *Carnegie Inst. Wash. Yearbook* (1933), p. 143.
3. H. W. Babcock, *Astrophys. J.* **118**, 387 (1953).
4. V. Ye. Stepanov and A. B. Severny, *Izv. Krymsk. Atrofiz. Observ.* **28**, 166 (1962); B. A. Ioshpa and V. N. Obridko, *Geomagnetizm i Aeronomiya* **2**, 541 (1962).
5. R. B. Leighton, *Astrophys. J.* **130**, 366 (1959).
6. V. Bjerknes, *ibid.* **64**, 93 (1926).
7. For a recent summary of our knowledge of sunspots, see R. J. Bray and R. E. Longhead, *Sunspots* (Chapman and Hall, London, 1964).
8. D. Osterbrock, *Astrophys. J.* **134**, 347 (1961).
9. V. Bumba and R. Howard, *ibid.*, in press.
10. R. B. Leighton, *ibid.* **140**, 1547 (1964).
11. ———, R. W. Noyes, G. W. Simon, *ibid.* **135**, 474 (1962).
12. N. R. Sheeley, thesis, California Institute of Technology (1965).
13. R. Kippenhahn and A. Schlüter, *Z. Astrofiz.* **43**, 36 (1957).
14. R. G. Athay, *Science* **143**, 1129 (1964).
15. H. W. Babcock and H. D. Babcock, *Astrophys. J.* **121**, 349 (1955).
16. H. D. Babcock, *ibid.* **130**, 364 (1959).
17. H. W. Babcock, *ibid.* **133**, 572 (1961).
18. V. Bumba and R. Howard, *ibid.*, in press.

Ant Venoms, Attractants, and Repellents

Secretions are used by ants in attack and defense and as chemical messengers in their social organization.

G. W. K. Cavill and Phyllis L. Robertson

Venoms, attractants, and repellents are glandular secretions which enable an animal to kill or paralyze the prey that forms its food or the food of its young; to convey messages to its fellows concerning food sources, mating, and the presence of its enemies; and to discourage or prevent those enemies from interfering with its social pattern.

General observations on insect secretions are by no means novel. But it is only in the last decade, with the development of more modern techniques in chromatography and spectroscopy, in bioassay and tissue culture, that research workers have acquired some of the analytical methods needed for fundamental studies on the small, even infinitesimal, amounts of these substances. In recent reviews, Beard (*1*) has discussed insect venoms and toxins, Roth and Eisner (*2*) have considered the defensive secretions or repellents of arthropods, and Karlson and Butenandt (*3*) have discussed the insect pheromones—that is, communication secretions. More recently, Wilson and Bossert (*4*) have considered chemical communication among animals. It now seems appropriate to make a critical examination of venoms, attractants, and repellents within one specific family, the Formicidae; to assess the present level of knowledge about their source, chemical constitution, and function; and, against this background, to consider possible lines for future development.

In using the terms *venom, attractant,* and *repellent* we are able to indicate the primary function of the secretion under consideration. But whatever its primary function, other, secondary functions may also be involved which are brought into play either deliberately or automatically. For example, a sting-bearing ant will normally use its venom offensively to kill the prey it hunts down for food (Fig. 1), but if the nest is attacked the ant may use its venom defensively to repel the intruder. Again, in the presence of an enemy, an ant may produce an alarm secretion which, at low level, may act as an attractant to members of its own nest, but, at higher concentration, may have the secondary function of stimulating them to aggressive or retreat behavior, while also repelling the foe (*4*). These are limitations to our basic classification of secretions as venoms, attractants, and repellents. The terms indicate primary function only, and convey no information concerning any secondary function the secretion may possess; they constitute broad, general categories and are not mutually exclusive.

Ant Relationships

In using ants as subjects for research on venoms, attractants, and repellents, we are dealing with a group of insects in which the production of these secretions has become of prime importance. Grounds for such an opinion are to be found in the evolutionary relationship of ants to other insects, in the characteristics of the order to which ants belong, and in the lines of specialization which have developed within the group itself.

Ants (family Formicidae) belong to the order Hymenoptera of the class Insecta, an order which has originated near the peak of the insect evolutionary tree (Fig. 2). Imms states (*5*), "If the Hymenoptera be judged by their behaviour, they must be regarded as including the highest members of their class. Structurally the majority of their species have attained an advanced degree of specialization which is only surpassed by the Diptera." In general, members of the Hymenoptera are notable for extreme mobility both on the ground and in the air, for the widespread adoption of parasitic modes of existence during development, for the use of the ovipositor as a sting, and for the evolution of social existence. In the family Formicidae, dependence on a functional sting reaches a maximum and is then lost, while social existence is developed to the greatest diversity and the greatest efficiency known in the Hymenoptera.

Within the Formicidae a remarkably wide range of habits is displayed (*6, 7*), which is significant in the present context because at least some habit differences can be correlated directly with changes from group to group in the rela-

Dr. Cavill is personal professor in organic chemistry, and Dr. Robertson is research fellow in biology, at the University of New South Wales, Kensington, Australia.