16 September 1960, Volume 132, Number 3429 SCIENCE

CURRENT PROBLEMS IN RESEARCH

Phenomena of the Solar Atmosphere

Astrophysicists are taking a new and more careful look at the outer layers of the sun.

R. Grant Athay

Astrophysicists long ago accepted the idea that because stars radiate energy into outer space, they cannot be in strict thermodynamic equilibrium. At the same time, however, they argued that the energy loss during a period of time characteristic of the collisional and radiative interactions was trivial in comparison to the energy density, and that, because of this, the stars could be considered to be in "local thermodynamic equilbrium," even in their surface layers. By this, they meant that Kirchoff's radiation laws were still applicable: the ratio of emissivity to absorptivity could be equated to the Planck function; absorption lines were formed in a cool stratum overlying a hotter surface; and, conversely, emission lines were formed in a relatively hot overlying stratum.

Parallel to this concept of local thermodynamic equilibrium were the postulates of radiative and hydrostatic equilibrium. According to these postulates, energy was transported outward by a radiative flux only, and the gaseous strata were supposedly quiescent.

The severest astrophysical test of these assumptions is found in the sun, which through its relative proximity allows a detailed study of its surface layers. Furthermore, the fortuitous phenomenon of a total eclipse of the sun provides a powerful means for studying the sun's external atmospheric layers. Two basic observations of the sun cast doubt on the foregoing postulates. On the one hand, the photospheric surface layers exhibit a granulation resembling convection cells, and on the other hand, the atmosphere overlying the photosphere extends to far greater heights than are compatible with the basic assumptions. Immediately above the photosphere and extending outward some 10,000 kilometers lies an irregular layer, known as the chromosphere because of its rosy-red color. This, in turn, gives way to a pearly-white corona often extending millions of miles into space.

Quite naturally, the first modification astrophysicists made in this classical picture, and the one most widely adhered to today, was postulation of the existence of a set of convection cells or "turbulent" eddies, which could explain both the granulation and the overly distended atmosphere. In the photosphere the convection cells have low velocity and do not seriously disturb the hydrostatic equilibrium. This is not so in the higher layers, however, and thus hydrostatic equilibrium has been dropped as an assumption. The assumptions of local thermodynamic equilibrium and radiative equilibrium are retained. The possibility that a coupling of the macroscopic motions to the microscopic degrees of freedom requires departures from these latter two assumptions is largely ignored, even though the macroscopic motions necessary to support the outer asmosphere are, of necessity, supersonic and might be expected to couple strongly with the microscopic degrees of freedom.

The startling discovery that resulted from the work of Grotrian and Edlén, that the mysterious emission lines in the spectrum of the corona are due to common elements in advanced stages of ionization, requiring temperatures of the order of a million degrees, forced a further reappraisal of the classical postulates. Both local thermodynamic equilibrium and radiative equilibrium were found to be invalid and were abandoned as totally unacceptable assumptions for the corona. Curiously, however, this new approach reestablished hydrostatic equilbrium as acceptable, because the high temperature was sufficient to explain the distended corona. Only recently have refined observations proved this assumption to be untenable. We are now forced to drop all three of the classical postulates.

Further down, in the chromosphere, the picture was not so clear. The assumptions of local thermodynamic equilibrium and of radiative equilibrium were retained with dogged determination. Supersonic turbulence was allowed to support the chromosphere but not to meddle with the internal degrees of freedom. Bit by bit this picture has crumbled as a result of improvements in eclipse data and the increasing evidence, growing mainly out of the work of R. N. Thomas and his co-workers, in support of large departures from local thermodynamic equilibrium in the chromosphere.

Solar physicists working with chromospheric problems almost universally recognize that such departures from the classical model of a stellar atmosphere exist. They are sharply divided on the degree and extent of the departures, however. Most of the standard refer-

The author is associate director of the High Altitude Observatory, University of Colorado, Boulder.

ence books in solar physics (*1*) adopt the point of view that departures from local thermodynamic equilibrium and radiative equilibrium are of secondary importance, whereas others, of a different school of thought (*2, 3*), consider them as being of primary importance. The evidence in favor of the latter group appears overwhelming at this time. Again, those who are willing to abandon entirely the assumptions of local thermodynamic equilibrium and radiative equilibrium find that hydrostatic equilibrium is restored in the low chromosphere and that the assumed supersonic turbulence is replaced by high kinetic temperature ($\approx 50{,}000°$) in a large part of the chromosphere (*4*).

Many astrophysicists, today, argue that these large departures from the classical stellar atmosphere pertain to the outer atmosphere only; that they do not influence the photospheric layers; and that they do not play a significant role in the formation of the Fraunhofer spectrum. On the answer to these questions hinge the important questions of the chemical abundances of the elements and of stellar models.

In the sun, the stronger Fraunhofer lines, such as the early Balmer lines and the H and K lines of Ca^{II} are known to be formed in the photosphere. If our current concept of chromospheric models (*2*) is correct, at least in principle, then these stronger lines are formed at kinetic temperatures well in excess of the photospheric temperature. Kirchoff's laws, as applied to these lines, are, therefore, totally invalid. Absorption lines can and do appear in a high-temperature atmosphere overlying a relatively cool radiating surface.

Recently, J.-C. Pecker (*5*) has demonstrated that marked departures from local thermodynamic equilibrium very probably exist in the fainter Fraunhofer lines of Ti^{II}, and that the accepted solar abundance of titanium is probably inaccurate by as much as a factor of 5. Analyses are in progress for other elements, which will undoubtedly yield new evidence on this important question. The work is tedious, but rewarding. It is inherently difficult, since the grossly simplifying assumptions of local thermodynamic equilibrium cannot be made. The most serious difficulty, however, is the lack of atomic optical and collisional cross sections for calculating rate coefficients.

Important new observational techniques are contributing to the rapidly changing concepts of solar phenomena. I shall mention them only briefly here and in the discussions that follow. The development of magnetographs for mapping magnetic fields in the solar surface layers, as pioneered by H. W. and H. D. Babcock, is a striking example of these new techniques. Radio astronomy is contributing heavily. Rocket observations of the solar ultraviolet spectrum, as pioneered by W. A. Rense, R. Tousey, and H. Friedman, are providing invaluable data that were not available to earlier generations. Balloon-borne telescopes flown by D. E. Blackwell and A. Dollfus in Europe and by M. Schwarzschild in this country are photographing the solar granulation in greater detail than ever before. The coronagraphs developed by B. Lyot for producing artificial eclipses continue to provide new insight on coronal and chromospheric problems.

Interpretation of Solar Activity

The solar physicist's concept and interpretation of the transient phenomena of solar activity has, to a large degree, paralleled his concept of the normal solar atmosphere. The initial interpretation of a dark feature on the solar disk was that it is relatively cool, and of a bright feature, that it is relatively hot. Now, one must analyze the problem more carefully before giving such an answer. For features observed in the continuous spectrum, the initial interpretation is undoubtedly the correct one. For those observed in the monochromatic light of the stronger Fraunhofer lines, it may be totally incorrect. A dark feature in a $H\alpha$ image of the sun, for example, is just as likely to be hot relative to the gases seen in the background radiation as it is to be cool. Similarly, a bright region, such as a flare, is not necessarily hotter than a dark region; it may simply be more dense.

These violations of Kirchoff's laws result, of course, from the failure of the conditions of thermodynamic equilibrium. Whether an object appears bright or dark in a given spectral line depends upon the balance between the various collisional and radiational transition rates influencing the number of electrons in the two energy levels involved in the formation of the line. Since the transition rate for collisions depends upon density as well as upon temperature, changes in either may affect the line formation. In thermodynamic equilibrium, this is not the case. Only the temperature may influence the radiation field.

Because of the complexity of the problem and the lack of adequate data, the solar physicist relies heavily upon the empirical approach as well as the analytical. As a case in point, consider the problems posed by solar flares.

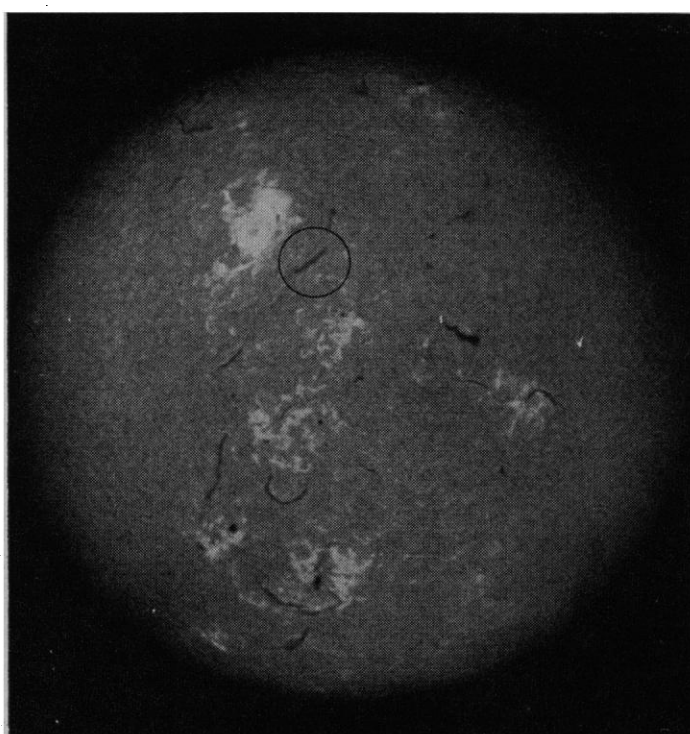
Flares are perhaps the most vivid manifestation of solar activity, both as directly observed on the sun and as indirectly observed through their associated terrestrial effects. They brighten explosively and often produce associated effects at remote points on the sun, as well as on the earth. Because of departures from local thermodynamic equilibrium in flares, the interpretation of the spectrum is difficult. A number of evidences suggest that a flare is of higher density than its surroundings. There is, however, no unambiguous evidence that it is also hotter. The temperature of flares that occur at coronal heights, which are not infrequent, is certainly lower than that of the surrounding corona. And it may well be that flares occurring in the chromosphere are also cooler than their surroundings. The question has not yet been answered, however, in anything approaching a satisfactory manner.

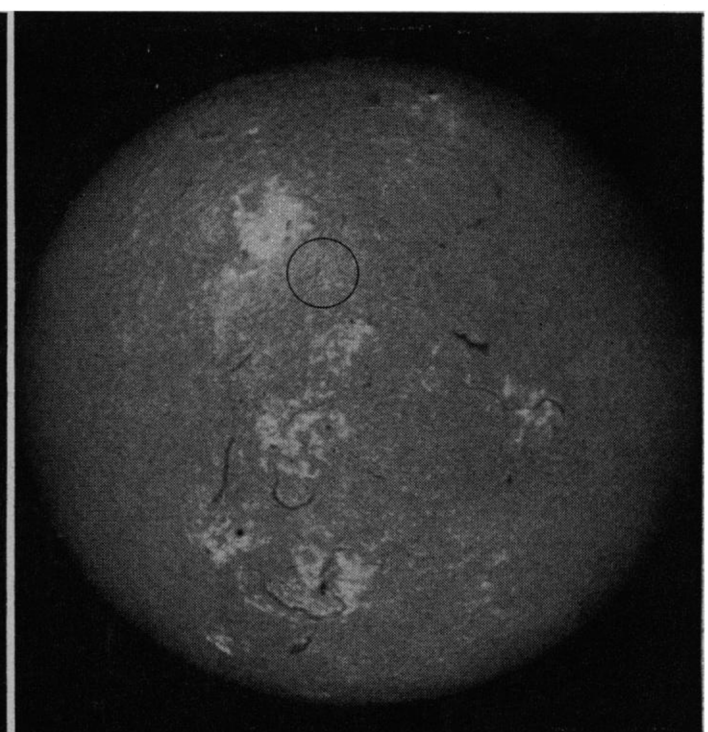
There is abundant empirical evidence that flares are often, if not always, accompanied by the ejection of fast streams of protons and electrons from the flare itself or from nearby regions. Evidence for this, at the solar end, comes from the observed ejection of matter in the form of surge and spray prominences, from associated bursts of radio noise, and from the associated observed effects at some distance from the flare, such as the momentary vanishing of prominences recently observed by Moreton (*6*) (see Fig. 1). The radio noise bursts are of different types, but each type is believed to result from localized macroscopic motions of the solar plasma. In the cases of the so-called "fast" and "slow" drift bursts, the motion of the source can be followed directly with high-resolution radio telescopes.

Flares are also accompanied by flashes of ultraviolet and x-radiation. The evidence for this comes almost entirely from their effects on the ionosphere but a few direct observations of the waning stages of flares from rockets by Friedman (*7*) and his co-workers have confirmed this conclusion. However, no direct observations are as yet available for the early or maximum phases of a flare, where the greatest changes are expected to occur.

If each flare were a repetition of every other flare, our problem would still be complex, but greatly simplified. Instead, each flare seems to be somewhat different. Some large flares produce ionospheric disturbances, geomagnetic storms, auroral displays, and increases in cosmic rays, all of major amplitude. Other flares that show similar characteristics in their visual spectrum produce small or undetected terrestrial effects. At the same time, what appear to be minor flares, from visual observation, may produce major terrestrial responses. These facts forcibly remind us that what we observe with our telescopes and spectrographs—for example, the brightening of certain spectral lines—is not related in a simple way to the physical processes in the flare phenomenon.

The great disadvantage in the empirical approach to the understanding of flare phenomena lies in its ambiguity. Does a flare itself eject clouds of matter from the sun? Does it emit intense bursts of ultraviolet and x-radiation? The answer to each of these questions might well be no. Instead, each of these separate manifestations of something unusual on the sun may be symptomatic of a larger, more widespread phenomenon that we have not





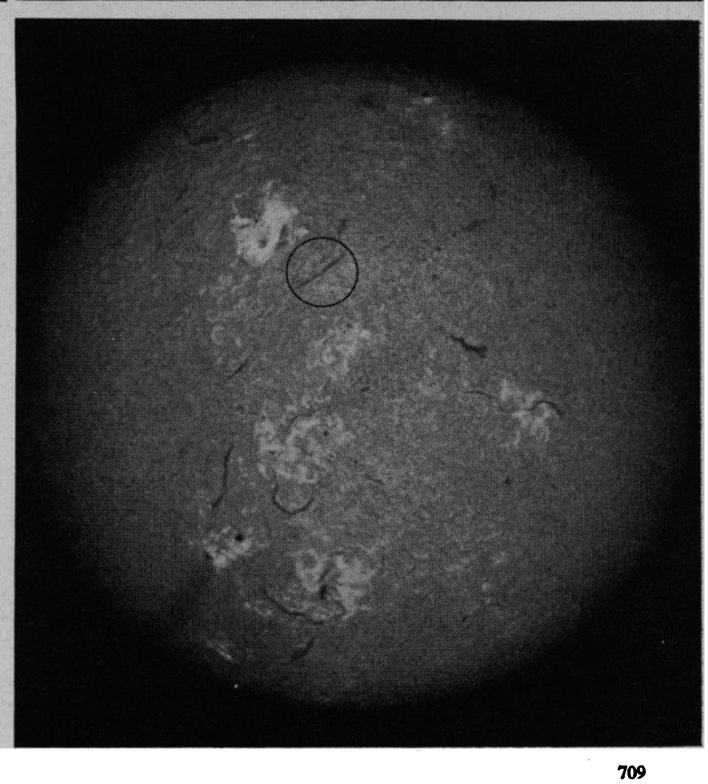


Fig. 1. A time sequence of photographs of the solar disk in a spectral band 0.5 angstrom wide centered on the Balmer-α line of hydrogen. The brightest region is a flare, and the dark, wormlike features are prominences lying usually 30,000 to 50,000 kilometers above the solar surface. The first picture (top left), taken at $21^h21^m30^s$ GMT, shows the flare during its period of rapid development and an encircled dark filament. The second (top right), taken at $21^h26^m40^s$ GMT, shows a more advanced stage of the flare near maximum brightness. Note that the filament has completely disappeared. The third picture (bottom right), taken only 1 minute 40 seconds later, shows the filament back in its original form. This momentary disappearance of the filament is believed to be a result of charge-exchange collisions between the neutral hydrogen in the filament and the protons of a proton-electron cloud ejected by the flare. [G. Moreton, Lockheed Solar Observatory]

yet observed in its broader aspects. There is some evidence that this is indeed the case. However, we cannot be certain until we understand the basic mechanisms involved in flare phenomena and the thermodynamic conditions within the flare itself. At present, we can do little more than speculate about the answers to these problems.

Indices of Solar Activity

The geophysicist in his search for the terrestrial consequences of solar activity turns to the solar physicist for a so-called "index" of solar activity, with the plea that it be "objective." What he actually wants is an index of corpuscular or short-wavelength electromagnetic radiation, which, of course, the solar physicist cannot supply "objectively" because he has no clear idea just what he should be objective about. To be sure, he can be objective about counting or measuring areas of sunspots or flares, but he cannot yet be objective in his use of such values, once they are obtained. Should he, for example, combine the area of a flare with its brightness, or would he get a better index by ignoring both of these aspects and taking the width of a spectral line as an index? The answer is simply that he doesn't know, nor will he know until the flare is understood as a physical phenomenon.

One may invert this reasoning and argue that we will not understand flares until we discover and understand the empirical associations between flares and other observable phenomena. To an extent this is true. However, it seems clear that the most straight-forward approach to the problem is to first arrive at a correct interpretation of the optical spectrum, which should at least furnish a thermodynamic description of the flare we are trying to explain. Once we have such a description, solution of the associated problems should be much simpler.

I do not intend by this discussion to discourage those who use solar indices in empirical studies of solar and solar-terrestrial phenomena. On the contrary, such studies are of great value and have proved to be fruitful. These studies are handicapped, however, by the unanswered questions concerning the basic meaning of solar indices. When physically meaningful indices, in terms of emitted solar radiation, are available, we can expect accelerated progress from sun-earth researches.

Origin of Solar Atmospheric Phenomena

The existence of the solar chromosphere and corona is a manifestation of a widespread departure from radiative equilibrium in the external layers of the sun. The origin and cause of this departure is as yet unspecified. Intuitively, one feels that it must, in some way, be associated with the hydrogen convection zone lying just below the photosphere. The outward decrease in temperature in the solar interior eventually reaches a value where protons and electrons start to recombine to form neutral hydrogen. When this occurs, the opacity increases sharply and the temperature gradient becomes steeper. As a result, the gases become convectively unstable and transport part of the energy flux by convection currents and part by radiation. It is presumably this convection that produces the photospheric granulation.

The convection zone very probably has side effects such as the generation and propagation of sound waves and hydromagnetic waves. Because of the outward decrease in density, certain of these waves tend to amplify and accelerate as they propagate outward. Eventually, they dissipate their energy into the propagating medium and effect a transfer of energy from the interior layers to the exterior layers by nonradiative processes. This energy, in turn, must be disposed of by the medium. Two questions stand out at this point in the theory. Where is the energy dissipated and how is it disposed of?

The last question is apparently the easier to answer. Most of the energy is disposed of by radiation, which can now escape into outer space because of the low opacity. The first question is not so readily answered. For many years we argued that because the temperature of the corona was about 1 million degrees, whereas the chromosphere was much cooler, the main energy dissipation must be in the corona. This is a fallacious argument. The greatest nonradiative dissipation of energy must occur where the greatest net loss of radiant energy occurs. Although we have not yet completely evaluated this net loss of radiant energy as a function of height in the solar atmosphere, it is clear that the maximum falls in the chromosphere rather than the corona. In this sense, the corona becomes secondary to the chromosphere. The high temperature in the corona results from the need to ionize the less abundant heavy elements several times before they can radiate efficiently enough to dispose of the energy dissipated therein. The chromosphere, on the other hand, disposes of its energy by radiation from the more abundant elements, hydrogen and helium, which radiate most efficiently at temperatures of some tens of thousands of degrees.

In the chromosphere we observe another phenomenon that probably plays a vital role in the energy cycle. This is the phenomenon of spicules, which are high-speed jets of gases ejecting upwards from the chromosphere into the corona (see Fig. 2). They occur more or less uniformly over the entire sun and number about 10^5 at any one time. Their diameter is of the order of 500 kilometers, and their outward elongation usually exceeds 10,000 kilometers during their few minutes of life. In a period of 24 hours, most of the chromospheric surface undergoes a spicule upheaval and the outward flux of matter is sufficient to replace the entire corona several times.

Spicules apparently originate in the low chromosphere, where the maximum net energy loss by radiation seems to occur, though we are somewhat uncertain as to the amount of this loss. The causes and effects of spicules are largely unknown. Nevertheless, it seems clear that they point to a basic instability in the chromospheric plasma. Their relationship to the energy balance of the chromosphere and corona presents one of the most challenging problems of solar physics.

Solar activity, like the chromosphere and the corona, seems to have its origin in departures from radiative equilibrium. The phenomena cluster around centers of activity marked in some stage of their evolution by sunspots. The active regions show much evidence of instabilities. These might well be of the same nature as those giving rise to spicules. One of the primary characteristics of active regions is abnormal magnetic field strength and, for many of the phenomena, abnormally high densities. Just how these characteristics arise and the role they play in the over-all phenomena are not clear. It seems likely, however, that solar activity owes its origin to the hydrogen convection zone, just as do the chromosphere and the corona. The acceleration of clouds and streams of charged particles from active regions seems inevitably related to their magnetic fields, but again the mechanisms are not understood.

Among the many unanswered ques-

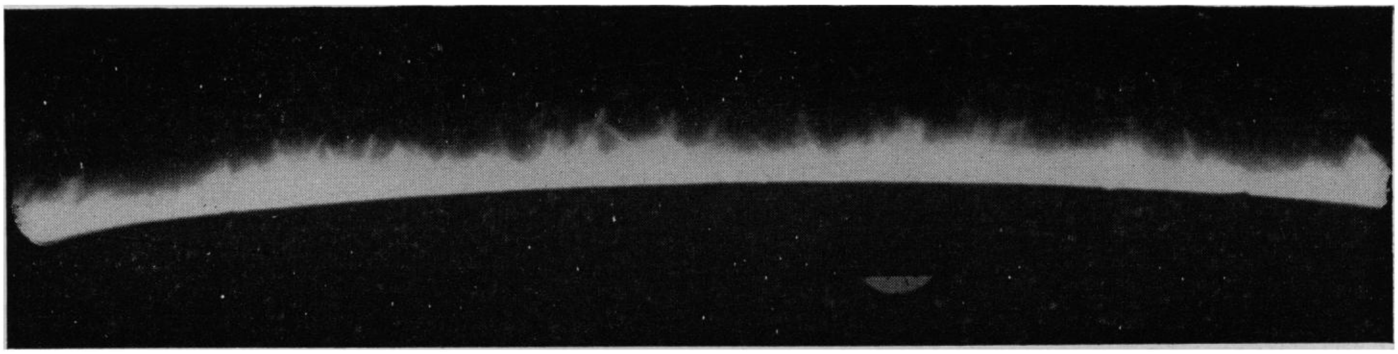


Fig. 2. A photograph of an artificial eclipse, showing the chromosphere at the extreme edge of the sun in the light of the Balmer-α line of hydrogen. The fine, bristle-like features are spicules. Their mean diameter is less than 1 second of arc, which is 725 kilometers at the distance of the sun. [R. B. Dunn, Sacramento Peak Observatory]

tions regarding phenomena of the solar atmosphere, those that stand out are: (i) What is the mechanism of energy supply? (ii) What is the role of magnetic fields? (iii) What is a flare? (iv) What are the instabilities that give rise to spicules and solar activity? (v) What are the nature and variability of solar corpuscular and short-wavelength radiation incident upon the earth as a result of solar activity?

New Horizons

Several recent developments have opened new horizons in solar and solar-terrestrial physics and promise to aid greatly in providing answers to some of the foregoing questions. Not the least of these is the recognition among solar physicists that they must consider departures from local thermodynamic equilibrium to be significant and that, as a result they need vastly more data than have been available in the past. These data are being gathered partly by extending the older techniques of observing and partly by developing new techniques. The most important of the new techniques are extraterrestrial observations from rockets and satellites and mapping of the magnetic fields in the surface layers of the sun. Among the extensions of older techniques are the high-speed photographic patrols of the solar disk, and of the chromosphere and corona at the limb, and the re-emphasis on eclipse observing to obtain detailed information on the spectrum of the upper photosphere and the lower chromosphere.

The extraterrestrial observations are furnishing immensely valuable information in the ultraviolet and x-ray regions of the electromagnetic spectrum. These observations will tell us much concerning the basic structure of the chromosphere and corona, provided we can obtain from laboratory experiments the necessary optical and collisional cross sections that are essential to a reasonable interpretation of the data. At these wavelengths the solar spectrum is characterized by a faint continuum crossed by numerous strong emission lines. A naive interpretation of these lines, following the concepts of local thermodynamic equilibrium, would indicate temperatures of about 7000° where the central core of the Lyman-α line of hydrogen is formed and about 20,000° where the corresponding line of ionized helium is formed. A more carefully considered treatment, however, indicates that the temperatures are actually of the order of 50,000° to 100,000° in these regions (*8*).

In the Lyman-α line of hydrogen the brightness of the solar surface is highly irregular, with most of the flux originating in active regions (*9*). Details of this sort are not yet available for other emissions in the short-wavelength regions of the spectrum, but there is little doubt that they will show similar concentrations in active regions. The total flux below about 1800 angstroms probably undergoes strong variations during the sunspot cycle and with the appearance of certain features of solar activity.

Corpuscular radiation from the sun is likewise being closely studied from satellites and from stratospheric balloons. A great deal has already been learned concerning this difficult problem, too much to summarize here. As in the case of the ultraviolet and x-ray spectrum, however, we are just at the threshold of challenging new discoveries.

In the meantime, much is being done from ground and from balloon-borne observatories. Observers in the United States, Russia, and England are carefully and painstakingly mapping the solar magnetic fields, hoping to establish their relationships to such phenomena as flares and other forms of solar activity. Observers the world over are constantly studying the solar disk, from high-altitude balloons, from mountain tops, and from sea level. Those with artificial eclipse-producing coronagraphs are trying to unravel the mysteries of the corona, of spicules, and of the upper chromosphere. Some (in many cases members of these same groups) are busy chasing eclipse shadows to remote regions of the earth, where, because of the eclipse produced outside of our terrestrial atmosphere, they find unparalleled observing conditions and opportunities for studying some of the vital solar problems.

Solar physics is making rapid advances at a constantly accelerating rate. Each recent decade has seen many revisions of concepts and many new data. The coming decade promises even greater progress.

References

1. G. P. Kuiper, *The Sun* (Univ. of Chicago Press, Chicago, 1953); C. de Jager, *Handbuch der Physik* (Springer, Berlin, 1959); L. E. Aller, *The Atmospheres of the Sun and Stars* (Ronald, New York, 1953); C. W. Allen, *Astrophysical Quantities* (Athlone, London, 1955).
2. R. N. Thomas and R. G. Athay, *Physics of the Chromosphere* (Interscience, New York, in press).
3. J.-C. Pecker and E. Schatzman, *Astrophysique Générale* (Masson, Paris, 1954).
4. R. G. Athay and H. R. Johnson, *Astrophys. J.*, **131** (Mar. 1960).
5. J.-C. Pecker, *Ann. astrophys.* **22**, 499 (1954).
6. G. Moreton, paper presented at the Cleveland meetings of the American Astronomical Society (1959).
7. H. Friedman, *Proc. I.R.E. (Inst. Radio Engrs.)* **47**, 272 (1959).
8. R. G. Athay, *Astrophys. J.* **131** (May 1960).
9. R. Tousey, *Sky and Telescope* **18**, 440 (1959).