13 March 1964, Volume 143, Number 3611 SCIENCE

The Challenge of Chromospheric Physics

Chromospheric plasma reveals complex interactions of magnetic, hydrodynamic, radiative, and thermal fields.

R. Grant Athay

The solar chromosphere is a chaotic region between the relatively cold, dense photosphere and the hot, tenuous corona. The great brilliance of the photosphere obscures the chromosphere under normal conditions, and the corona, because of its great dimensions and spectacular beauty, dominates the image of the sun at total solar eclipse. The chromosphere is visible for but a few seconds during total eclipse, and it is natural that it has been largely ignored in favor of the more spectacular corona.

In the yesterday of solar physics the chromosphere was regarded as little more than a puzzling curiosity. Although it possessed features that challenged the imagination, it was not considered of primary importance in the overall picture of the solar atmosphere. Today the importance of the chromosphere can hardly be overemphasized in solar physics, and efforts to properly interpret chromospheric phenomena are leading to concepts that may have repercussions throughout astrophysics.

The chromospheric emission-line spectrum was known by early observers of eclipses to include lines of both neutral and ionized helium, neither of which were to be found among the Fraunhofer absorption lines in the spectrum of the solar disk. Spectral lines of singly ionized metal atoms in the chromosphere were known to be enhanced in strength relative to lines of their parent neutral atoms. Hence it was customary to describe the chromospheric spectrum qualitatively as resembling a laboratory "spark" spectrum, and the photospheric spectrum as resembling a laboratory "arc" spectrum.

The higher state of excitation of the chromospheric spectrum relative to the photospheric spectrum could, supposedly, be explained by the lower density of the chromospheric gases. This is the familiar "pressure ionization" effect noted in laboratory experiments and predicted by theory. It was clearly recognized, however, that the pressure effect could not excite helium atoms because of their high excitation potential. The answer to this puzzling feature was sought in extrachromospheric sources of energy.

Strengths of helium lines were known to increase, relative to other lines, in the higher layers of the chromosphere bordering the corona. The corona was known to be very hot and capable of producing strong ultraviolet radiation. It was suggested that this radiation penetrated the chromosphere and induced fluorescent radiation from helium atoms. So little was known about either the corona or the chromosphere, when this suggestion was put forward, that adequate quantitative tests were not feasible.

Theories of stellar interiors and atmospheres were well advanced in the early decades of this century. One prediction of theory was that the temperature of the solar atmosphere should decrease monotonically outward, approaching a "boundary temperature" of about 4700°K. The chromosphere was in the predicted boundary region, and its temperature was assumed to be the boundary temperature. The corona overhead, with a temperature near one million degrees, was clearly recognized as an anomaly in this picture. The corona was so obviously different from the chromosphere, however, that it was perhaps logical to assume that whatever produced the corona had no direct influence on the temperature of the chromosphere.

Warning Signs

This picture of the chromosphere as an essentially passive boundary layer with a few spectral peculiarities induced by the corona persisted until the 1940's. Many years earlier, however, the warning signs had been posted, some by a few individuals, such as Menzel, who were skeptical of the theory and some by observers studying the geometrical properties of the chromosphere.

Eclipse observers reported early in this century that the chromosphere extended some 12,000 kilometers beyond the visible solar limb. The boundary-layer theory predicted a chromosphere no more than 200 kilometers in depth. Hence this vast extent of the chromosphere presented a sharply defined, unexplained problem to the astrophysicist. The density of matter decreased much more slowly with height in the chromosphere than it did in the photosphere. One of the main questions concerning the chromosphere, therefore, was, What supports this greatly distended atmosphere?

Various support mechanisms were

The author is professor and head of the department of physics at the University of Utah, Salt Lake City. This article was written in part while Dr. Athay was at the High Altitude Observatory, Boulder, Colo.

postulated and rejected until McCrea proposed in 1929 that macroscopic turbulent motion supported the chromosphere by exerting a pressure similar to that exerted by thermal motions of atoms and molecules. The velocity of turbulent motion required to produce sufficient pressure was estimated to be about 16 kilometers per second, provided the temperature was near 4700°K throughout the chromosphere. Subsequent observations of widths of chromospheric helium lines indicated that the helium atoms do indeed have random velocities of about 16 kilometers per second, and the theory of support by turbulent eddies was supposedly strengthened. Thus, the helium lines in the chromospheric spectrum, which were themselves an anomaly, were closely tied to two separate and distinct facets of chromospheric phenomena: (i) they were excited by external radiation from the corona, and (ii) they were broadened by the same turbulent motions that, supposedly, supported the chromosphere against the strong solar gravity.

The upper layers of the chromosphere were carefully scrutinized by early observers of the sun. Secchi, in 1877, described the upper layers as having the visual appearance of a "burning prairie." The important point to be drawn from Secchi's description is that the upper boundary of the chromosphere is irregular and constantly changing. In 1941 Roberts discovered that the small flame-like protuberances observed by Secchi (which Roberts named spicules) have mean lifetimes of the order of 5 to 10 minutes, that they grow vertically from small bumps on the chromospheric boundary to columnar features some 10,000 to 15,000 kilometers long at speeds of the order of 25 kilometers per second, and that they are rather uniformly distributed over the entire sun. This spicule structure of the upper chromosphere presented a new challenge to the solar physicist. Figure 1 is an excellent photograph of the spicule detail observed in a small segment of the solar limb.

A Closer Look

The usual theory of stellar atmospheres is based on three basic assumptions—those of hydrostatic equilibrium, radiative equilibrium, and thermodynamic equilibrium. According to the first, the atmospheric gases are essentially quiescent, and there are no large-scale velocity fields that seriously perturb the atmosphere. According to the second, the energy flow through the atmosphere is carried by the radiation field. Both conduction and convection of energy are negligible. The third assumption requires that every microscopic process occurring in the gas be balanced by its inverse process. These assumptions make it possible to construct models of stellar atmospheres with a minimum of observational data. Conversely, if enough data are available, all three of these assumptions become subject to testing. In considering the chromosphere, the assumption of hydrostatic equilibrium had been quickly abandoned in favor of an assumption of low temperature and support by turbulent eddies. The other two assumptions, however, were not so readily challenged.

What may be termed a rebirth in chromospheric physics started along several different lines. In 1941 Redman suggested that the widths of chromospheric lines could best be interpreted in terms of a temperature of 35,000°K. This suggestion of high temperature simultaneously challenged the concept that the chromosphere was a simple boundary layer in radiative equilibrium and made it largely unnecessary to postulate a separate support mechanism; support would be provided by the purely thermal energy of the gas.

As a result of Redman's suggestion, Thomas in 1948 and Giovanelli in 1949 made calculations showing that the hydrogen spectrum of the chromosphere was perhaps as compatible with a temperature of 35,000°K as with a temperature of 4700°K, provided the assumption of thermodynamic equilibrium was dropped. The basis of Redman's suggestion has since been shown to have been faulty, and the initial calculations of Thomas and Giovanelli are now recognized to have been oversimplified and inadequate. Nevertheless, the basic assumptions of radiative and thermodynamic equilibrium had been challenged and had not been exonerated by available data.

One of the most significant effects of these challenges was the inspiration they provided for observing the chromospheric spectrum in much more detail than had been previously achieved. Natural eclipses of the sun still provide the best opportunities to study the chromospheric spectrum in detail. (Figure 2 is a print of a portion of one of the chromospheric spectrograms obtained at the 1962 eclipse.) However, the chromosphere is visible for only a few seconds during an eclipse, and eclipses are not ideal for studying the geometrical properties of the chromosphere. For this purpose one turns to coronagraphs and other solar telescopes with which observations can be made of the chromosphere either at the limb of the sun at times other than eclipse or as seen in projection against the disk of the sun. Using this approach, we are able to get better spatial resolution of chromospheric features and to study their time evolution. Figure 1 illustrates the structural detail observable at the limb, and Fig. 3 and the cover illustrate the structure observed on the disk.

In the last few years important additional sources of chromospheric data have become available. The sun's observable spectrum extends beyond the infrared into radio wavelengths. At millimeter and centimeter wavelengths the observed radiation arises primarily in the chromosphere. Large radio telescopes are constantly scanning the sun at these wavelengths. The data obtained are extremely valuable, but the full capabilities of this new technique have not yet been exploited. Rockets and satellites carrying solar telescopes and spectrographs above the earth's ionosphere are mapping out and monitoring ultraviolet and x-radiation from the sun. The most intense radiations in this previously inaccessible region of the solar spectrum are now known to originate in the chromosphere rather than the corona. This one fact, apart from all others, shows clearly and unequivocally that the chromosphere is of major importance in the solar atmosphere.

Rocket and satellite observations, like radio observations, are still in their infancy, and the data leave much to be desired. Nevertheless, these techniques are already playing an important role in our studies of the chromosphere. They will undoubtedly play greater roles in the future.

All these factors and events have contributed to the current vitality of chromospheric physics. There are other important factors that I will explain after discussing the chromosphere itself.

Chromospheric Geometry

Any interpretation of the chromospheric spectrum in terms of thermodynamic properties of the gas depends heavily upon the particular geometry assumed. For example, the intensity of the spectral line emitted by a gas is, in some instances, proportional to the square of the gas density. Thus, a feature of the spectrum with a given intensity may be explained either by assuming a uniform distribution of matter in the sun's atmosphere at density ρ or by assuming the occurrence of discrete features, with, say, a density of 10 ρ, occupying 1 percent of the volume.

Observations of chromospheric structure are relatively easy to obtain but difficult to interpret. Structure observed against the disk by isolating selected narrow bands of the solar spectrum that are formed in the chromosphere is rich in detail (see Figs. 3 and 4). So far, however, efforts to interpret this detail have been largely unsuccessful. No satisfactory location of the observed features in terms of a vertical height scale has been achieved.

Observations of structure at the limb remove the ambiguity in height but present other difficulties. For example, let us imagine that the sun is an opaque sphere bristling with spicules, much like a pincushion bristling with pins. When we observe this model sun from some distance away, it appears as a flat, opaque disk projected on the plane of the sky, with the spicules extending beyond the circumference. The circumference of the solar disk, which is the great circle girdling the spherical sun in the plane of the sky, is referred to as the solar limb. Spicules extending beyond the solar limb may intersect the solar surface near the limb or they may intersect it some distance away. From simple geometrical considerations, it follows that any feature, such as a spicule, that protrudes more than 1000 kilometers vertically out of the opaque solar surface from a point located within 35,000 kilometers of the limb would extend beyond the limb when projected on the plane of the sky. If such features covered 1 percent of the sun's surface and were distributed uniformly, a line of sight tangential to the opaque solar surface would intercept five of them. If the features were similar in appearance, they would blend into an unresolved mass indistinguishable from a homogeneous atmosphere.

It follows, then, that any observed set of distinct features at the solar limb that are similar to each other in appearance, that occur more or less uniformly over the sun, and that average more than 1000 kilometers in vertical extent probably cover less than 1 percent of the sun's surface area. Conversely, discrete features in the solar atmosphere are not likely to be detected at the limb by direct observation if they are small and numerous

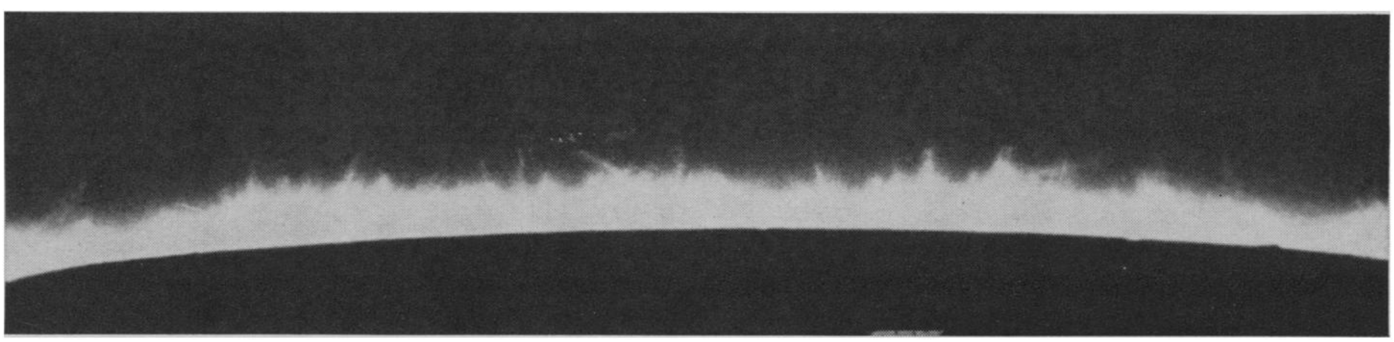

Fig. 1. Spicules in the upper and middle chromosphere, as observed in the Hα line of hydrogen. [Courtesy R. B. Dunn, Sacramento Peak Observatory, Air Force Cambridge Research Laboratories]

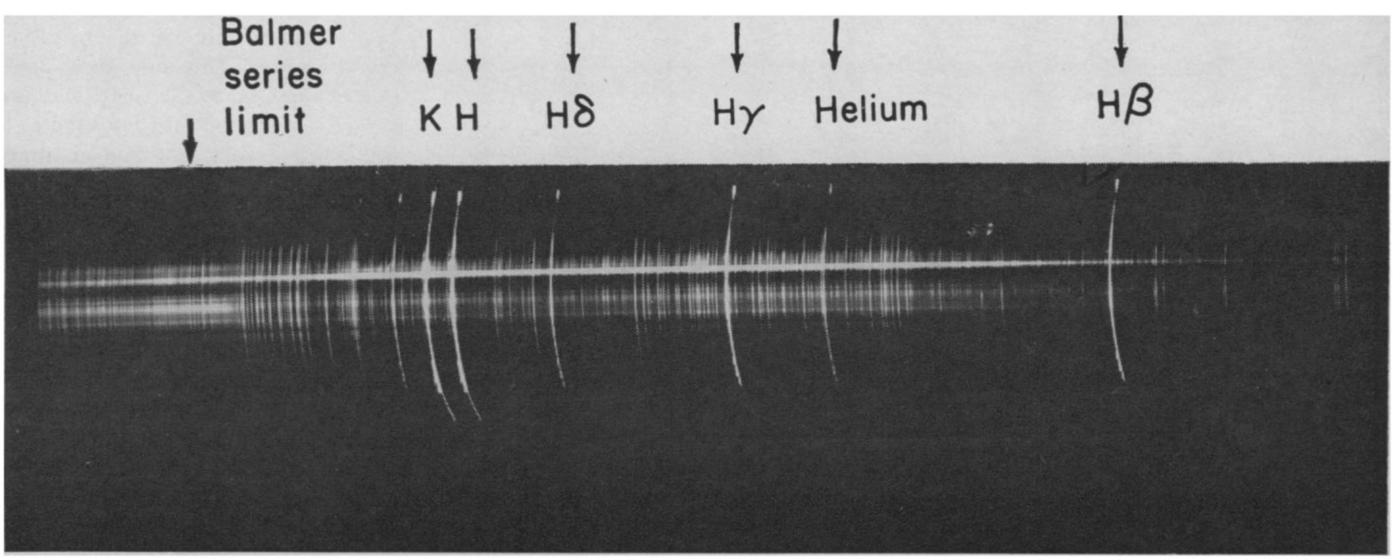

Fig. 2. The chromospheric blue-violet spectrum observed, in Lae, New Guinea, at the 1962 total eclipse by a joint expedition of the High Altitude Observatory, Sacramento Peak Observatory, and the National Bureau of Standards.

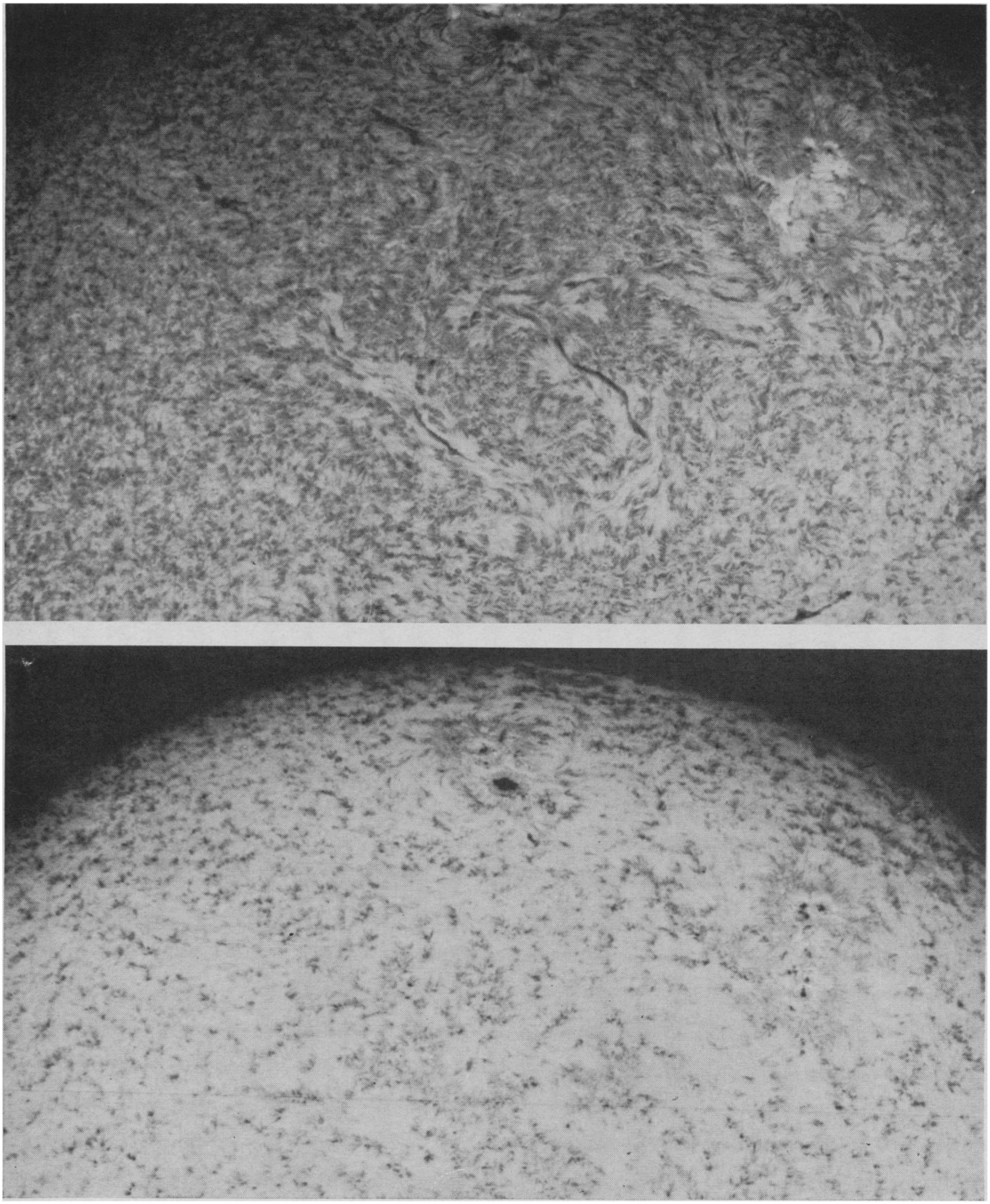

Fig. 3. Spectroheliograms of the solar disk, showing the chromospheric structure as observed (top) at +0.35 angstrom from the center of the Hα line and (bottom) at +0.70 angstrom from the center of Hα (see, also, cover, which shows the chromospheric structure in the center of the Hα line). Note the characteristic changes in the appearance of bright plages near the sunspots and of the long dark filaments, as well as the finer mottling on the disk. The cover picture, Fig. 3 (top), and Fig. 3 (bottom) represent progressively decreasing heights in the chromosphere. [Courtesy R. B. Leighton, California Institute of Technology and Mt. Wilson Observatory]

enough to cover more than 1 percent of the solar surface.

Spicule structure is observable at the limb, at times other than eclipse. Detailed counts of spicules observed in the Hα line of the hydrogen Balmer series are available for all heights above 3000 kilometers. At the 3000-kilometer level spicules cover only about 0.3 percent of the chromospheric surface (*1*). On the basis of observations made during the 1958 eclipse, Suemoto (*2*) claims that spicules are resolved as low as 1500 kilometers in the chromosphere.

At heights where we see spicules clearly resolved, essentially all of the chromospheric emission comes from the spicules. Thus, the entire line spectrum of the chromosphere above about 3000 kilometers, and possibly above 1500 kilometers, essentially arises from spicules covering less than 1 percent of the sun's surface at any one time. The regions between the spicules are evidently very much less dense, and they have an entirely different type of line spectrum. It appears, in fact, that as low as 3000 kilometers the interspicule regions are coronal in character —that is, they have a temperature near a million degrees.

A faint continuum spectrum produced by scattering of photospheric radiation by free electrons is observed at heights where spicules occur. Since this feature of the spectrum depends linearly on the density, instead of the square of the density, it is likely that regions between spicules contribute substantially to it. The "mean" electron density at a height of 6000 kilometers above the solar limb, as determined from continuum data, is about $8 \cdot 10^9$ electrons per cubic centimeter (*1*). Spectral-line data give an electron density within spicules of about 10^{12} electrons per cubic centimeter (*3*), and the spicules cover about 0.001 of the surface. Thus, if all the continuum were produced by spicules, the "mean" density would be $1 \cdot 10^9$. It seems clear, then, that the interspicular regions contribute most of the continuum spectrum obtained for this height.

Emission in the helium lines is faint for the lowest layers of the chromosphere, rises sharply to a maximum for the region at about 1100 kilometers, and then decreases slowly with increasing height. Analysis of the data indicates that for heights below about 3000 kilometers the helium emission does not arise in the same volume elements of the chromosphere as the rest of the lines in the visual spectrum (*3*). The magnesium spectrum further indicates that at 1000 kilometers, and possibly at 750 kilometers, there are already sharp departures from spherical symmetry in the chromosphere (*4*).

Monochromatic observations of the solar disk show unmistakably that structure exists in the solar atmosphere at all heights where the Fraunhofer spectrum is formed. This structure is evidenced, as in Figs. 3 and 4, by fluctuations in brightness commonly re-

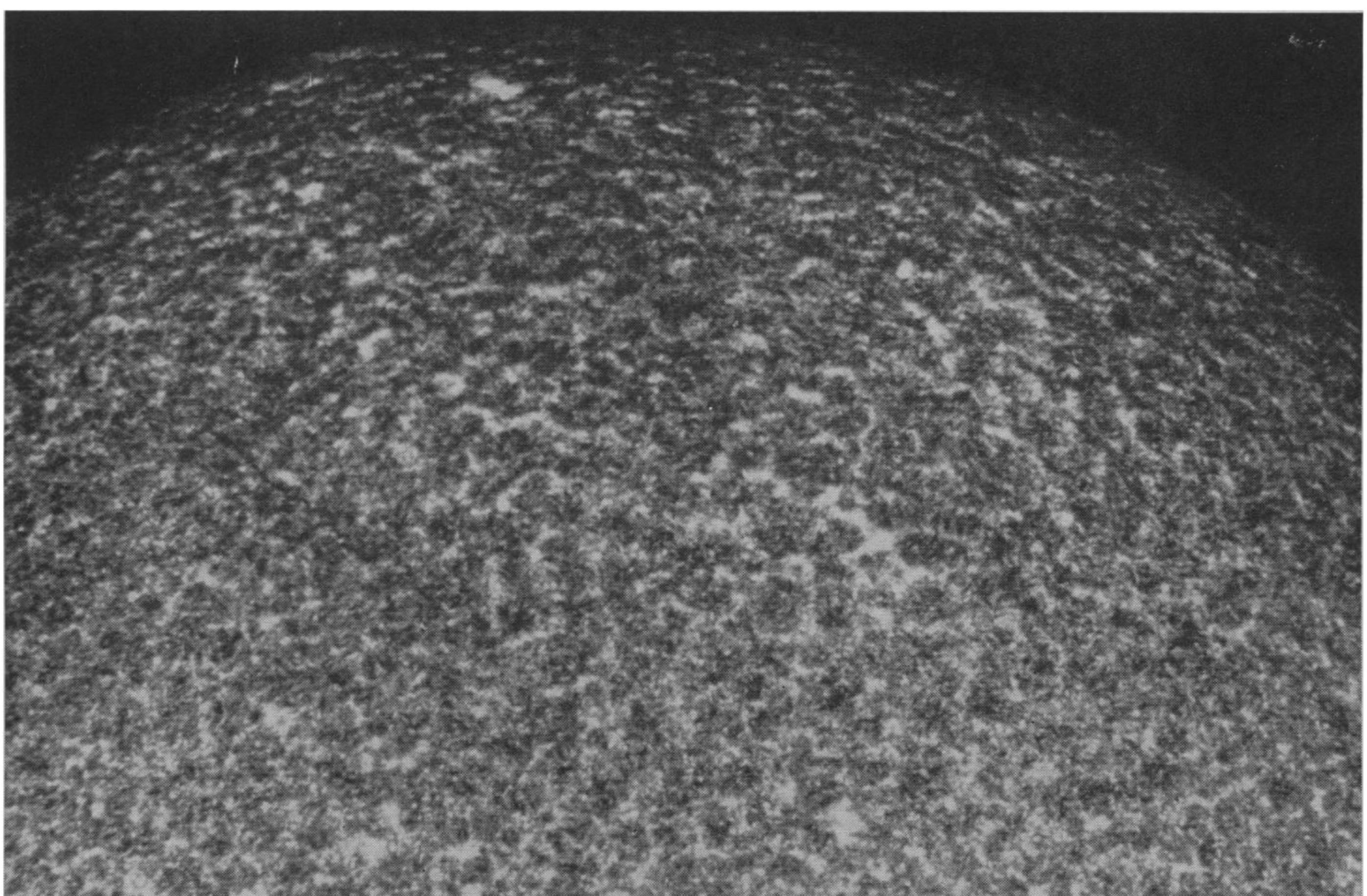

Fig. 4. Chromospheric structure as observed at −0.18 angstrom from the center of the K line of ionized calcium. Note the bright network pattern as well as the smaller features. [Courtesy R. B. Leighton, California Institute of Technology and Mt. Wilson Observatory]

ferred to as mottling. Interpreting the mottling is difficult. Since there is structure, the "surface" observed at a given wavelength must have "hills" and "valleys" that are directly related in some manner to the mottling. As yet we have devised no satisfactory way of determining depths of the "valleys" or heights of the "hills." Partly for this reason, and partly for others that I note later, we have not succeeded in interpreting the mottling in terms of fluctuations in density and temperature in the atmosphere. We do know that the centers of some of the faint Fraunhofer lines arise from levels deep in the photosphere and that the centers of such strong lines as the H and K lines of ionized calcium, the Hα line of hydrogen, and the b_1 line of neutral magnesium arise, generally, from regions well above the 1000-kilometer level in the chromosphere. Since these lines, as well as lines of faint or intermediate strength, show fluctuations in brightness from point to point on the solar surface, structure exists throughout the chromosphere and photosphere.

Velocity Fields

McCrea's suggestion of a chromosphere supported by turbulence becomes meaningless when considered in terms of present concepts of chromospheric geometry, particularly the spicule structure. Spicules are short-lived geyser-like ejections rising out of the low chromosphere at speeds of 25 to 30 kilometers per second. At maximum elongation spicules appear to "pause," maintaining their maximum height for about 30 percent of their lifetime; then they either fade from view or retract along their original paths. This life history sounds suspiciously like that of a projectile moving under the influence of gravity. The mean speeds, the maximum heights, and the percentage of lifetime near maximum height are consistent with such an assumption. However, we have not been able to detect systematic accelerations in spicule motion, and we have no reason to suppose that whatever forced the spicule upward in the first place ceased to act before we first observed the spicule.

Polar spicules are oriented predominantly in radial directions with respect to the sun, whereas equatorial spicules tend to be inclined to the radial direction. In regions of the solar limb of the order of 100,000 kilometers in length, spicules seem to move in well-defined patterns suggestive of large-scale force fields, resembling magnetic-field patterns observed in photospheric layers.

At any one time there are about 10^5 spicules on the sun in some stage of evolution. At a given point on the sun a spicule ejection occurs, on the average, about once a day. However, there is growing evidence that spicules have preferential locations on the sun, at least over periods that are long as compared to the life of a spicule.

Spicules have a density of about 10^{-12} gram per cubic centimeter, which is essentially constant with height. They are believed to originate near the 1000-kilometer level, where the mean atmospheric density is of this same order. The upward flow of mass in spicules is sufficient to replace the entire corona in about an hour's time and is much more than is required to maintain the solar wind. The flux of kinetic energy from spicules at the 5000-kilometer level is comparable to the rate of energy loss from the corona by radiation. Therefore, it appears that spicules may play a major role in the dynamics and energetics of the corona.

One of the more interesting features of the solar velocity field is the recently discovered vertical oscillatory motions (*5*). The lower layers of the chromosphere and the upper and middle photosphere are composed of numerous cell-like features about 3000 kilometers in diameter. These cells oscillate up and down about their average position in the atmosphere with periods of about 300 seconds and mean velocities of 0.5 to 0.8 kilometer per second. Both the period and the mean velocity of oscillation appear to decrease as height in the atmosphere increases. A given cell oscillates through two or three cycles, then remains relatively quiet for about an hour.

The nature of the waves in the solar atmosphere giving rise to the oscillatory motions is of immediate interest. The initial upward displacement in an oscillating cell begins first in the lower layers of the cell and moves progressively upward, the motion of the upper layers lagging behind the motion of the lower layers. After about one full oscillation, however, the entire cell apparently moves in phase. The complexities of these phase lags and the period of the motion suggest that the wave motion is complex and that to describe the waves as, for example, either simple sound waves or gravity waves is not adequate.

Somewhat higher up in the chromosphere large-scale cells of the order of 30,000 kilometers in diameter are observed against the solar disk. These cells, shown in Fig. 4, are relatively long-lived and exhibit a flow of matter parallel to the sun's surface and outward from the centers of the cells at a velocity of about 0.4 kilometer per second. Near the edges of the cells, and somewhat deeper in the atmosphere, there is a slow downward streaming of matter. These same border regions, however, seem to be the favored locations for spicules (*6*).

In the same regions of the chromosphere and photosphere where the oscillatory cells and the larger long-lived cells are observed there is yet another kind of macroscopic motion. Widths of spectral lines formed in these layers indicate randomly moving, macroscopic "cells" of gas too small to be observed as discrete entities but with velocities considerably larger than those of the observed cells. In the photosphere these random velocities are about 2 kilometers per second. They increase in the chromosphere to about 5 kilometers per second at heights where the largest cell structures are observed. Thus, the oscillatory motions and the steadier motions observed in the large cells are superposed on smaller-scale motions of considerably greater velocity. No unifying description of the total velocity field has yet been attempted.

The low chromosphere is known to depart from spherical symmetry and to possess both systematic and random macroscopic velocity fields. However, none of the macroscopic structures, except spicules, have velocities as high as the thermal velocities of protons and hydrogen atoms that make up the bulk of the atmosphere. It therefore makes some sense to check the assumption of hydrostatic equilibrium. Temperatures and densities derived from continuum data in which spherical symmetry is assumed are in fact consistent with the assumption of hydrostatic equilibrium up to heights of about 1000 kilometers, where the consistency breaks down rather sharply. We interpret this to mean that geometrical irregularities and macroscopic motions at heights below 1000 kilometers are in the nature of perturbations in a spherically symmetric atmosphere, not the gross departures from symmetry represented by the spicules.

Magnetic Fields

Direct evidence of magnetic fields in the chromosphere is difficult to obtain. The presence of strong magnetic fields in the photosphere and in prominences elevated above the chromosphere clearly indicates, however, that magnetic fields permeate the chromosphere. The homogeneity of spicule motions suggests that they are guided by magnetic lines of force. The kinetic-energy density in spicules is about 5 ergs per cubic centimeter. Hence magnetic fields strong enough to control the spicule motion would have to be of the order of 10 gauss or greater.

The large cells of about 30,000-kilometer diameter referred to in the foregoing section exhibit a relationship to magnetic-field patterns. Cells of this type are most easily observed by isolating, spectroscopically, the wavelength bands just off the centers of the H and K lines of ionized calcium (see Fig. 4). When observed in this way, the cells appear as relatively dark areas bordered by brighter vein-like features. These bright border regions form a mosaic pattern on the solar disk, referred to as the calcium network.

Magnetic fields are detected in the solar atmosphere by measuring the Zeeman splitting of spectral lines. Observations of the Zeeman splitting are most easily made in the fainter, narrower Fraunhofer lines that are formed in photospheric layers. Such observations show enhanced magnetic-field strengths in a pattern corresponding to the calcium network (*7*), but at a lower level in the atmosphere. As I have previously noted, spicules, which are in still higher layers than the network, seem to occur preferentially in these same locations.

The correlation between enhanced magnetic-field strength in the photosphere and the bright calcium network in the chromosphere extends to other features of the chromosphere, such as plages, that produce brighter-than-average radiation in the H and K lines of ionized calcium. The same is true for chromospheric regions that produce enhanced brightness in the hydrogen lines.

In the case of the calcium network the increased chromospheric brightness is associated with hydrodynamic motions within the large-scale cells. If the network regions are indeed the favored locations for spicules, as has been suggested, the association of the network regions with regions of enhanced magnetic-field strength takes on added importance. The formulation of a physical theory linking chromospheric brightness, hydrodynamic motions, and magnetic-field strength has not been seriously attempted as yet.

Thermodynamic Equilibrium

The brightness of the continuum spectrum of the chromosphere was reliably measured for the first time at the 1952 solar eclipse. In the lowest layers of the chromosphere, several different sources of continuum emission are of importance. These are: (i) recombination emission from the negative hydrogen ion, (ii) recombination emission from hydrogen, (iii) scattering of photospheric radiation by free electrons, and (iv) scattering by neutral hydrogen atoms. (Recombination emission occurs when a free electron loses energy by combining with an atom or ion, in this case a hydrogen atom or a proton.)

The base of the chromosphere is set arbitrarily to coincide with the limb of the solar disk, as observed at a wavelength of 5000 angstroms. The limb marks the height in the solar atmosphere at which the solar gases have an optical thickness of unity with respect to a tangential ray. The opacity of the chromosphere decreases rapidly with height, and above about 100 kilometers we may consider the chromosphere a transparent medium that does not absorb in the continuum even when the ray is tangential. This greatly simplifies analysis of the continuum data.

The spectrum of the low chromosphere is rich in spectral lines. In many of these lines the chromospheric opacity is low, as it is in the continuum. However, for many of the stronger lines the chromosphere has high opacity, which complicates analysis of the line spectrum.

Because the chromosphere has low opacity at some wavelengths, it is bathed in the strong radiation field of the photosphere; at the same time, it is unable to retain its own radiation. Thus the radiation field within the chromosphere is not necessarily in equilibrium with the thermal properties of the chromosphere. In particular, if there is an energy source other than the radiation field that raises the kinetic temperature of the chromospheric gases, the disequilibrium between the radiation field and the kinetic temperature may become very pronounced.

In such a situation the usual simplifying conditions of thermodynamic equilibrium do not apply, and all of the relevant microscopic rate processes must be carefully considered in analyzing the spectrum. The thermodynamic state of the gas can no longer be completely described by two thermodynamic variables, such as density and temperature. The concept of a temperature loses its universal meaning, and separate "temperatures" may be required to describe such properties as energies of free particles, relative populations of bound energy states, and intensities of radiation fields.

Even in the case of those spectral lines that are strongly modified by the chromosphere, there may be large differences between the "temperature" characterizing the intensity of radiation in the line and the "temperature" characterizing the kinetic energy of free particles. In the energy equations for the radiation field there are diffusion terms and source terms. The diffusion terms represent the effect of the radiation field in the line as it is built up through the entire atmosphere. The source terms include both direct collisional effects and the effects of radiation and rates of collision at other frequencies. These latter effects may be dominated by radiation fields originating completely outside the local environment, and therefore they will not necessarily bear a simple relationship to such parameters as the local kinetic temperature.

As an example, let us consider the Lyman-α line of hydrogen at $\lambda 1216$ (1216 angstroms) as it is formed in the chromosphere. The intensity of the radiation in this line depends upon the relative populations of the first two bound energy levels of hydrogen. In thermodynamic equilibrium these relative populations can be described solely by the kinetic temperature of the gas. In fact, however, the population of the second energy level of hydrogen in the chromosphere depends heavily upon the intensity of the radiation in the Balmer continuum near $\lambda 3640$, which is controlled almost exclusively by the photosphere and hardly at all by the local kinetic temperature of the chromosphere. Similarly, the population of the first energy level depends heavily upon the intensity of the radiation in the Lyman continuum near $\lambda 912$, which originates much deeper in the chromosphere than the center of the Lyman-α line does. Thus, in the chromosphere the Lyman-α radiation field is partially controlled by completely external influences and does not mimic the thermodynamic character of its local environment. In regions of the chromosphere where the center of the Lyman-α line is formed, the kinetic temperature of the gas is of the order of 70,000°K (*8*). The escaping radiation, however, is characteristic of radiation from a black body at about 7000°K, which is over a million-fold less intense than radiation one would expect from a black body at 70,000°K. Similar effects, some much less pronounced and some more so, are represented in almost all chromospheric lines.

The radiation field of a gaseous atmosphere such as the chromosphere can be predicted only by solving the energy equations simultaneously with equations specifying the populations of energy level in terms of all the relevant parameters. In most cases both the energy equations and the population equations for several lines must be solved simultaneously in order to get a valid solution to any one of the equations. The coupled equations are complex and have been solved analytically only through recourse to simplifying assumptions of questionable validity.

Conversely, the observed spectrum of a gas such as the chromosphere cannot be properly interpreted without the same detailed analysis. It follows also that, to derive all the thermodynamic properties of the gas, one would need many more details of the spectrum than are needed in analyzing an atmosphere in thermodynamic equilibrium.

Thermodynamic Structure

The complexity involved in interpreting the chromospheric spectrum has severely curtailed attempts to understand the thermodynamic structure of the chromosphere. This difficulty is compounded by complex geometry that affects both the distribution of radiation in space and the exact boundary conditions to be applied in solving the energy equations. Nevertheless, progress is being made, and many interesting details of chromospheric thermodynamics are emerging. Almost without exception these details are controversial. The following discussion represents my own views and those of some of my colleagues. Unless otherwise specified, I use the word *temperature* to mean the kinetic temperature.

For the chromosphere below about 1000 kilometers, one may reasonably hope that the assumption of spherical symmetry provides a satisfactory first approximation. Furthermore, the continuum emission is quite sensitive to the local kinetic temperature and density. Effects of departures from thermodynamic equilibrium are reflected in continuum data, but they do not dominate the analysis. Thus, in all respects, the region below 1000 kilometers is the region of the chromosphere that is simplest to analyze.

Above a height of 1000 kilometers in the chromosphere, emission in the continuum in the visual spectrum becomes rather insensitive to the local kinetic temperature because of the predominance of electron scattering, and we obtain information mainly on the mean density of the atmosphere. Information on the temperature and geometrical structure up to heights of 2500 kilometers comes mainly from helium-line data in combination with Balmer-continuum data. For heights of 3000 kilometers and more, we rely mainly on helium- and hydrogen-line data obtained at eclipse, and on the shapes of line profiles for individual spicules.

Helium-line data in the visual spectrum place strong boundary conditions on the temperature structure. The precise interpretation of these data, however, involves the solution of energy equations for the first two lines of the resonance series lying in the far ultraviolet. Specifically, the intensities of the singlet lines relative to one another and to their corresponding triplet lines depends primarily upon gradients in energy flux in the first two resonance lines (*3*). These gradients depend upon the opacity of the chromosphere, and the opacity depends, in turn, upon temperature and density. It is relatively easy, in this case, to relate opacity to temperature and density but very difficult to relate flux gradients to opacity. This latter problem is under careful study by Jefferies (*9*) and others.

Temperature decreases outward through the photosphere. Near the top of the photosphere, which is also the base of the chromosphere, the temperature reaches a relatively sharp minimum of about 4500°K. It rises to about 6200° at 500 kilometers and to about 7200° at 1000 kilometers. The density at the base of the chromosphere is about $2 \cdot 10^{-8}$ gram per cubic centimeter. At 500 kilometers it has dropped to about 10^{-10} gram per cubic centimeter, and at 1000 kilometers, to about $4 \cdot 10^{-12}$ gram per cubic centimeter. To a fairly good approximation, this region of the atmosphere is in hydrostatic equilibrium (*1*).

Near the 1000- to 1200-kilometer level, helium lines appear quite suddenly in the spectrum and reach their maximum intensity. Helium emission and emission from hydrogen and metals can be shown to originate in distinctly different volume elements. Spherical symmetry can no longer be assumed as a first approximation, and the chromosphere takes on a spicule-like character. Essentially, all of the line spectrum characteristic of neutral and singly ionized atoms arises in the spicules or spicule-like features.

Within spicules the mean density is about 10^{-12} gram per cubic centimeter, and it is relatively constant with height. The relatively few emission lines of elements heavier than oxygen found in the visual spectrum for heights above 3000 kilometers arise in a cold spicule core whose temperature is about 10,000° to 12,000°K. Helium lines arise mainly in a hot envelope surrounding the cold spicule core at temperatures near 50,000°K. The relative proportion of hot envelope to cold core increases with increasing height. Above 3000 kilometers the hot envelope dominates the spicule structure, and at lower heights the reverse is true. Hydrogen and oxygen emission lines arise in both the cold core and the hot envelope. Below 3000 kilometers they arise mainly in the cold cores because the cores predominate, and above 3000 kilometers they arise mainly in the hot envelopes, which predominate at those heights.

At heights of about 1100 kilometers in regions where no spicules are present, temperature rises rather abruptly to about 50,000°K and density drops to about 10^{-14} gram per cubic centimeter (*3*). In the layers immediately above there appears to be an approximation to hydrostatic equilibrium. Somewhat higher, but below 3000 kilometers, temperature rises abruptly to a value characteristic of the corona (about 10^6 deg K) and density drops to about 10^{-15} gram per cubic centimeter.

The hot-envelope model is the simplest that has been constructed for the spicule from data presently available. Surrounding the hot spicule envelope is a still hotter corona, and we may visualize the hot spicule envelope as an intermediate region between the hot corona and the cold spicule core.

This schematic model is characterized by three distinct temperature regimes. These regimes have been predicted theoretically as well as inferred from observational data (*1*). They exist because of the chemical composition of the sun and because the chromosphere and the corona lose energy mainly by radiation, whereas they gain it mainly by nonradiative processes. Two conditions relative to energy balance must be satisfied in the solar atmosphere. The gas must be capable of radiating away the energy input, and the temperature field must be stable against perturbations in the energy input. That is, perturbations must be damped rather than amplified.

Hydrogen is the major chemical constituent of the sun. In the temperature regime below about 12,000°K, hydrogen radiates efficiently and perturbations in energy input are damped by offsetting changes in radiative output. Above this temperature hydrogen becomes increasingly less efficient as a radiator and energy perturbations are amplified rather than damped. Less abundant constituents in the solar atmosphere may become more important than hydrogen as radiators of energy if they lose some of their electrons. The maximum radiating efficiency of a given atomic species increases approximately in proportion to $(z + 1)^4$, where z is the excess nuclear charge. Helium is about 0.1 as abundant as hydrogen in the solar atmosphere, and when singly ionized $(z + 1 = 2)$ it is capable of radiating more energy than hydrogen. Similarly, many of the heavy elements, which are about 10^{-4} as abundant as hydrogen, could, under suitable conditions, radiate more energy than hydrogen if they were to lose ten or more electrons.

The temperature structure of the solar atmosphere depends, of course, on the nature of the energy-input mechanism as well as on the energy-loss mechanism. If the energy input permits, as it apparently does in the solar atmosphere, the temperature of the atmosphere may rise until the less abundant elements lose enough electrons to radiate with about the same efficiency as the maximum radiating efficiency of hydrogen.

For each ionization stage of each element the radiating efficiency first increases with increasing temperature, passes through a maximum, then decreases as temperature continues to increase. In the temperature regime where the radiating efficiency of the ion producing the dominant energy loss from the solar atmosphere is increasing with increasing temperature, the temperature of the atmosphere tends to be stable against perturbations in energy input. When the radiating efficiency is decreasing with increasing temperature, the temperature of the atmosphere tends to be unstable against perturbations in energy input. Hence, we postulate that the temperature in the solar atmosphere may have only those values at which the radiating efficiency of the dominant radiator increases as temperature increases. If we further postulate that the radiation loss per unit mass must remain at a value near the maximum radiation loss provided by the hydrogen atoms in the unit mass, we conclude that the outward increase of temperature in the solar atmosphere will occur in three step-like plateaus rather than gradually. The three plateaus correspond to the stable radiation regimes of hydrogen, singly ionized helium, and heavy elements that have lost ten or more electrons.

The three temperature regimes in the chromosphere and corona inferred from spectroscopic data do, in fact, provide conditions such that either neutral hydrogen, or helium with one electron removed, or a heavy element with about ten electrons removed is the dominant radiator and has not yet reached its maximum radiating efficiency. Hydrogen radiation predominates in the low chromosphere and in spicule cores where the temperature is below about 12,000°K. Radiation from singly ionized helium predominates in the upper chromosphere and in spicule envelopes where the temperature is about 50,000°K. Radiation from heavy elements, such as silicon, magnesium, and iron, predominates in the corona.

The source of energy and the mechanism of energy transfer for the chromosphere and corona remain a mystery. The sharply rising temperature in these regions signifies that the dominant energy input is a nonradiative process which we may classify broadly as "mechanical" in nature. The only identifiable sources of energy capable of supplying the amounts estimated for the chromosphere are the energies contained in the thermal and convective motions in the denser, deeper layers of the photosphere.

The general magnetic field of the sun has an energy density of only about 0.1 erg per cubic centimeter. The energy loss in Lyman-α radiation alone, when averaged over the chromosphere, would completely consume the energy of the magnetic field about once each minute. Hence, the magnetic field cannot be the main source of chromospheric energy. On the other hand, the magnetic field may provide a vital link in the transfer of energy from the deeper solar layers into the chromosphere.

The energy transported in the photosphere by convection processes is of relatively little importance in comparison to the intense radiation of the photosphere, but it is more than is needed to maintain the chromosphere and corona. The most promising of the mechanisms thus far proposed for transferal of the convective energy to thermal energy involve the generation of pressure waves that have a strong acoustic component in the photosphere. As these waves propagate upward they evolve into shock waves and hydromagnetic waves. The pressure shock waves dissipate their energy mainly in the lower chromosphere, and the hydromagnetic waves propagate on into the upper chromosphere and the corona (*10*). Until we understand the energy requirements of the chromosphere in considerably more detail than we do now, however, the details of the energy-input mechanism must remain unsettled.

Only a part of the radiant energy of the chromosphere need be supplied by mechanical sources. The remaining, and possibly dominant, portion is photospheric radiant energy that is scattered by chromospheric atoms and only appears to be arising within the chromosphere. Determination of the energy budget of the chromosphere, therefore, requires a precise separation of the scattered and the generated energy. This poses a difficult problem in radiative transfer, as well as requiring reliable knowledge of chromospheric geometry and thermodynamics. Such problems are in the forefront of current chromospheric research.

Chromospheric Activity and Our Space Environment

Many of the phenomena of solar activity, together with their pronounced terrestrial effects and their influences on our local space environment, arise within the chromosphere. Most flares, plages, and surge and spray prominences, and some radio bursts at wavelengths shorter than about 10 centimeters, fall in this category. Chromospheric flares and flare-associated phenomena represent the most violent forms of solar activity and produce profound effects in our outer atmosphere and in our space environment.

Bursts of far-ultraviolet and x-radiation that produce ionospheric storms, great proton showers that bombard the polar regions of the earth, high-energy solar cosmic rays, and the lower-energy plasma clouds that result in geomagnetic storms and produce auroral displays are all strongly associated with flares. The immediate source of some of these irregular solar radiations is probably the corona surrounding the flare, or perhaps even the photosphere below the flare. Nevertheless, the chromosphere is intimately involved in the flare phenomenon.

The most intense emission lines and continua in the far-ultraviolet region of the solar spectrum, as observed by rockets and satellites, arise from chromospheric hydrogen, helium, and heavier elements in early stages of ionization. These radiations produce and control much of the terrestrial ionosphere and play an important role in the energy budget of the earth's outer atmosphere.

An understanding of the sun's atmosphere and of solar activity in its many forms is of great importance in all our space programs and in upper-atmosphere research. Strong additional interest is being generated by current emphasis on the production and study of high-energy plasmas in attempts to produce controlled nuclear fusion in the laboratory. Many of the phenomena observed in the laboratory have a marked resemblance to phenomena of the solar atmosphere. Thus, we now have, for the first time, the possibility of performing laboratory experiments in chromospheric and coronal physics. Similarly, plasma physicists are recognizing that the chromosphere and the corona are "laboratories" in which interesting experiments are continuously in progress. Interaction of the laboratory studies of plasma physics with studies of chromospheric and coronal physics will undoubtedly increase our understanding of this complex and interesting region of the solar atmosphere.

References

1. R. N. Thomas and R. G. Athay, *Physics of the Chromosphere* (Interscience, New York, 1961).
2. Z. Suemoto, in *Transactions of the International Astronomical Union*, vol. 11B (Academic Press, New York, 1962), p. 206.
3. R. G. Athay, *Astrophys. J.* **137**, 931 (1963).
4. ——— and L. L. House, *ibid.* **135**, 500 (1962).
5. R. Leighton, R. Noyes, G. Simon, *Astrophys. J.* **135**, 474 (1962); J. W. Evans and R. Michard, *ibid.* **136**, 493 (1962).
6. G. Simon, thesis, California Institute of Technology (1963).
7. R. Leighton, *Astrophys. J.* **130**, 366 (1959).
8. D. C. Morton and K. G. Widing, *ibid.* **133**, 596 (1961).
9. J. T. Jefferies, *ibid.* **132**, 775 (1960).
10. D. Osterbrock, *ibid.* **134**, 347 (1961).

Microscopic Brains

The behavior of insects and vertebrates may not differ qualitatively to the extent that had been supposed.

V. G. Dethier

The study of animal behavior is unique among the sciences because it begins historically and methodologically with human behavior, prescinds from human experience, and projects this experience into other animals. It is thus more disposed to subjectivity and introspection than the other sciences and constantly labors under the burden of containing these biases within the bounds of their historical context. The study of man himself is further complicated by the fact that the investigator is trying essentially to understand himself, and others through himself, and in so doing is employing a brain to understand a brain.

Students of behavior tend to seek in other animals that which they believe exists in themselves. They look for motivation, drive, emotion, perception, consciousness, ideation, mood, sensation, and learning. Common sense assures us that it would be absurd to deny the existence of these phenomena. Those to whom an appeal to common sense borders on scientific heresy need only peruse the *Handbook of Physiology* and dwell upon the chapter headings: "Drive and motivation," "Emotional behavior," "Attention, consciousness, sleep, and wakefulness," "Perception," "Thinking, imagery, and memory." These are real phenomena. Faced with defining them, however, we bog down in a morass of ignorance, confusion, anthropomorphism, and verbal gymnastics to escape anthropomorphism.

Given this background, how can one ever study these states outside the context of human behavior? Certainly the most obvious and tangible approach is a search for physiological correlates. When a dog which is teased by a stranger bares its fangs, raises its hackles, snarls, and lunges, we say that it is enraged. Whether it is or not we shall probably never know, any more than we can ever know when a fellow human is enraged. On the other hand, we can ask meaningful and testable physiological questions about the dog's behavior in this situation, which so closely mimics our own emotion in comparable situations. We can investigate the conditions under which hair is erected, adrenalin is secreted, teeth are bared. It may even be possible to gain some insight into possible affective components of this behavior through employment of the self-stimulation techniques discovered by Olds and Milner (*1*).

Viewed in this light, it is clear that higher animals, mammals especially, exhibit a rich repertoire of behavior comprehended under the terms motivational, emotional, and so on which seems to be absent to varying degrees in the so-called lower animals. The farther removed an animal is from ourselves, the less sympathetic we are in ascribing to it those components of behavior that we know in ourselves.

The author is professor of zoology and psychology at the University of Pennsylvania, Philadelphia. This article is adapted from a lecture delivered 28 December 1963 at the Cleveland meeting of the AAAS.