Peter V. Foukal

Solar Astrophysics

Second, Revised Edition

WILEY-VCH Verlag GmbH & Co. KGaA

This Page Intentionally Left Blank

Peter V. Fouka

Solar Astrophysics

This Page Intentionally Left Blank

Peter V. Foukal

Solar Astrophysics

Second, Revised Edition

WILEY-VCH

WILEY-VCH Verlag GmbH & Co. KGaA

Author

Dr. Peter V. Foukal
Heliophysics, Inc.
pfoukal@world.std.com

Cover Picture The sunspots are part of NOAA Active Region 10030 observed on 15 July 2002 using the Swedish 1-m Solar Telescope. Kindly provided by G. Scharmer

Library of Congress Card No. applied for

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library.

Bibliographic information published by

Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

Printed in the Federal Republic of Germany.

Printed on acid-free paper.

Typesetting Kühn & Weyh, Satz und Medien, Freiburg
Printing betz-Druck GmbH, Darmstadt
Bookbinding Großbuchbinderei J. Schäffer GmbH & Co. KG, Grünstadt

ISBN 3-527-40374-4

To Lizzie

This Page Intentionally Left Blank

Preface to the Second Edition

Understanding of the sun has advanced in several important areas since the first edition of this text in 1990. Nobel prize-winning neutrino measurements have confirmed recently that we really do understand how the sun and stars shine. Analyses of solar irradiance measurements have shown how the luminosity of the sun and similar stars is modulated by the surprisingly effective valve of photospheric magnetic activity. Helioseismic measurements of the sun's radial temperature profile now agree with the standard model of the solar interior to better than ½%. They also reveal the sun's internal rotation profile, showing an interesting shear layer between the radiative core and convection zone, and confirming earlier evidence for another, immediately below the photosphere.

Improved simulations of solar convection now reproduce photospheric granulation remarkably well, with prediction of the larger observed supergranular convection scale as the next aim. When rotational influences are included, such models can explain the tendency to angular momentum conservation that seems to be observed in supergranular-scale overturning. This modest but encouraging success may offer a clue toward more ambitious explanation of the sun's deeper differential rotation with latitude. Dynamo models based on flux transport by the observed motions of the solar interior, including new evidence for a meridional circulation, can reproduce increasingly detailed aspects of the solar magnetic cycle. But a successful dynamical model remains a goal.

In the corona, new observations of dramatically anisotropic temperature distributions, and of large abundance anomalies relative to photospheric layers, encourage new thinking about the mechanisms most likely to heat and accelerate the solar wind. Measurements from the *Ulysses* spacecraft reveal the 3-D structure of the heliosphere and indicate that the fast wind from coronal holes is the prevalent flow except at low latitudes, and around solar activity maximum. New data from the SOHO mission and other spacecraft have helped to understand the important role that coronal mass ejections (CME's) play in geomagnetic disturbances and solar energetic particle events ("space weather"). But key processes of e.g. solar cosmic ray acceleration seem to originate within the flare site itself. The hottest plasmas around the flare source are now being imaged spectrally in hard X-rays and gamma rays from the RHESSI mission, and using radio frequency timing techniques.

The aim of this book remains the same – to provide an introductory but quantitative discussion of solar phenomena and relevant astrophysical processes. The need

Solar Astrophysics. Peter V. Foukal

ISBN: 3-527-40374-4

for this seems to increase as the research literature becomes ever more specialized and challenging to assimilate (especially for veterans like the author!). Happily, some obsolete material could be deleted, making room for those new developments that seem most significant for this introductory account. The references have also been updated, but some classics have been retained; they sometimes provide the clearest explanations. Several new figures have been included to illustrate specific advances in understanding, or without apology, simply to show readers spectacular new images. Hannes Alfven was motivated by films of solar prominences to formulate his Nobel prize-winning theory of conducting fluid dynamics, and the beauty of the sun's magnetic structure always provides an attraction to solar research.

I am grateful to numerous colleagues for their advice on recent research developments- the opportunity to renew such contacts has provided one of the main satisfactions of revising this text. I am particularly indebted to those who contributed figures and tables, in some cases produced specifically for this book. Except in Chapter 1, I emphasize results over techniques. This is not intended to overlook the debt that solar research owes to the many skilled scientists who devote their careers to development of sophisticated ground based and space-borne instrumentation. Some detailed accounts of the interesting instruments,spacecraft and ground based facilities used in solar observations are cited in references, particularly in Chaper 1.

Study of the sun has provided me with great personal enjoyment since 1969, when I first arrived at CalTech to work as a post doctoral fellow with Hal Zirin. Previously, my graduate research in galactic astronomy in Manchester, England with Z. Kopal and J. Meaburn, had taken me to the Pic du Midi Observatory in the French Pyrenees, where I was impressed by my first views of the sun – some through Bernard Lyot's famous coronagraph. Hal's knowledge of the sun, and his generosity and enthusiasm provided a supportive environment for my first optical and radio measurements of the sun.

An opportunity to move to Harvard in 1972 provided my first experience with (then novel) "big" space science, through participation in the SkyLab mission. The Harvard EUV data provided an exciting new window on the sun's atmosphere, and I benefited from many discussions on this topic, and on my other interests in solar rotation, luminosity variation, and plasma electric field diagnostics, with Gene Avrett, Andrea Dupree, David Layzer, Bob Noyes, and other CFA colleagues. A prominent visitor at Harvard at that time was Jack Eddy, whose investigations of solar activity and climate records provided all of us with a new outlook on sun-climate studies. Jack's mentoring after I left Harvard to continue my solar studies through private research firms, opened opportunities for which I continue to be grateful.

My unusual career path in solar research would not have been possible without the multiplicity and pragmatism of funding sources that is a great strength of US science. Internationally, solar research continues to enjoy ample support for wonderful ground-based and space-borne facilities. Its future depends on attracting imaginative and skilled young scientists to achieve the potential of these facilities. If this book helps to interest a few such individuals in solar research, it will have met its goal.

Peter Foukal
Nahant, Massachusetts

Contents

Solar Astrophysics. Peter V. Foukal

ISBN: 3-527-40374-4

1
Development of the Ideas and Instruments of Modern Solar Research

1.1
Early Telescopic Discoveries on the Sun

Scientific inquiry into the nature of the sun began around 1610 with the first telescopic observations of sunspots (called maculae or blemishes) by J. Fabricius, Galileo Galilei, C. Scheiner, and others in Western Europe. Naked-eye sightings of dark markings on the sun's disk had been reported in China and elsewhere at least 15 centuries earlier, and this evidence was almost certainly familiar to Galileo and his contemporaries. But the telescope showed that the westward motion of the spots was fastest near disk center and relatively slower near the east and west limbs. Using this observation, Galileo was able to argue from the projection effects expected in circular motion that the spots must be dark markings on the sun's surface rather than planets or other bodies in distant orbit and transiting its disk.

By 1630, the Jesuit astronomer Scheiner had used the spots' daily motion to accurately measure the sun's equatorial rotation period of 27 days as seen from the orbiting earth and to determine the 7° tilt of its equator relative to the ecliptic plane. Scheiner also detected the longer rotation period of spots at high solar latitudes and inferred that the sun's angular rotation rate decreases toward the poles. The well-known confrontation between Galileo's presentation of these early discoveries on the sun and the Church's doctrine of immaculate and stationary celestial bodies made a direct impact on the cosmology of the day that remains unique in the history of solar research.

To understand the fate of solar research in the two centuries following these promising beginnings, we should recall that the sun's nature was about as dimly perceived in Galileo's time as is the nature of quasars today. A useful estimate of its distance and size became available only after 1673, when the distance of Mars from Earth was measured using that planet's parallax observed between Paris and French Guiana. The Earth-sun distance then followed from J. Kepler's third law (of 1619) to within 10 % of the correct value. The sun's mass relative to that of the Earth (determined from the orbital periods and distances of the Earth and moon) was not published by Isaac Newton until some 50 years after Galileo's first observations. The Earth itself was properly "weighed" by H. Cavendish's experiment only in 1798. Even the order of magnitude of the sun's surface temperature remained uncertain

Solar Astrophysics. Peter V. Foukal

ISBN: 3-527-40374-4

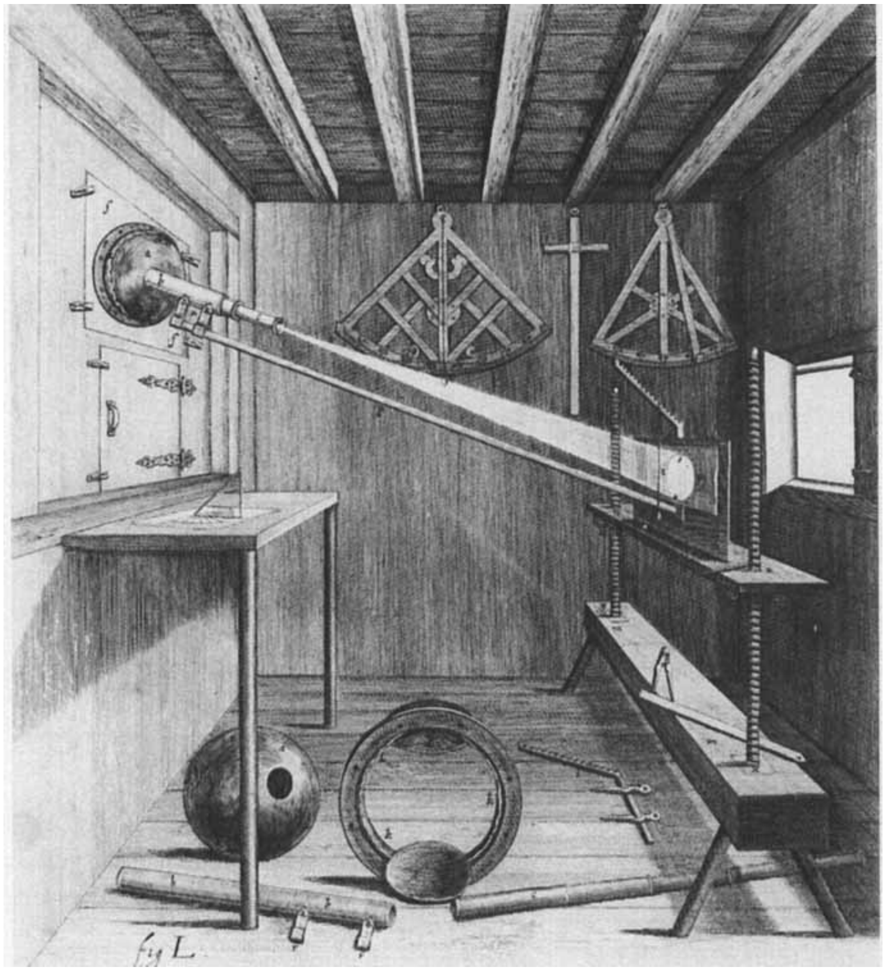

Fig. 1-1 Eyepiece projection scheme used by J. Hevelius to observe sunspots and faculae. From his *Selenografia*.
By permission of the Houghton Library, Harvard University.

until after the 1880s, when the fourth-power relationship between the temperature of an opaque body and its emittance came to be accepted through the experimental work of J. Stefan and the thermodynamic arguments set forth by L. Boltzmann.

Given the slow development of the physics and chemistry required to interpret solar observations, it is not surprising that for roughly two centuries after the first telescopic discoveries solar research consisted mainly of sunspot observations. Observers in the early 17th century used non-achromatic refractors with spatial resolution of 10 to 15 arcsec, and for detailed work the solar image was usually projected onto a screen (Fig. 1-1). (Never look at the sun through any optical instrument without expert advice – serious eye damage could result). Drawings by J. Hevelius (Fig. 1-2) and Scheiner show that the darkest, roundish central umbra of the spot is often surrounded by a roughly annular lighter region, the penumbra. The relative positions and areas of individual spots within a group were observed to change from hour to hour.

These observers also noted bright irregular patches comparable in size to spots, that could often be observed in association with sunspots near the limb. Scheiner called them faculae or little torches. The contrast of faculae is much lower than that of spots, and they are not visible at all in white light near disk center. Both spots and faculae can often be observed to recur on subsequent solar rotations, implying lifetimes of up to several months. Figure 1-2 shows the march of faculae and a sunspot across the disk from day to day.

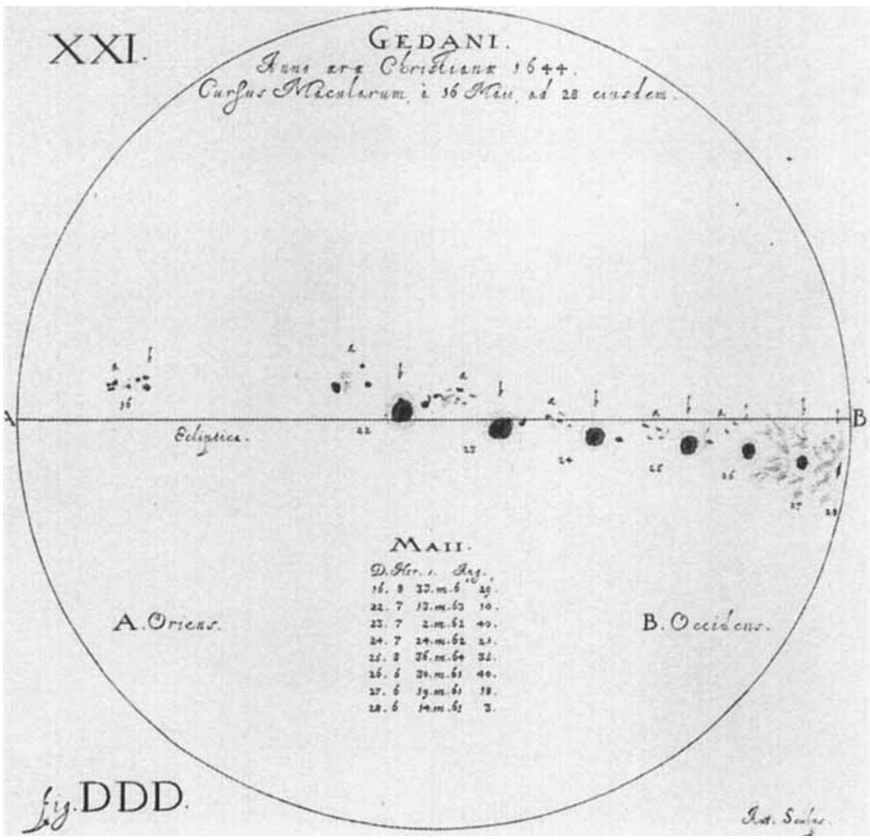

Fig. 1-2 Sunspots and faculae observed by J. Hevelius on May 16–28, 1644. The position of the large spot is shown on successive days as solar rotation moves it from east to west across the disk. Faculae (hatched) are seen only near the limb. From the *Selenografia*. By permission of the Houghton Library, Harvard University.

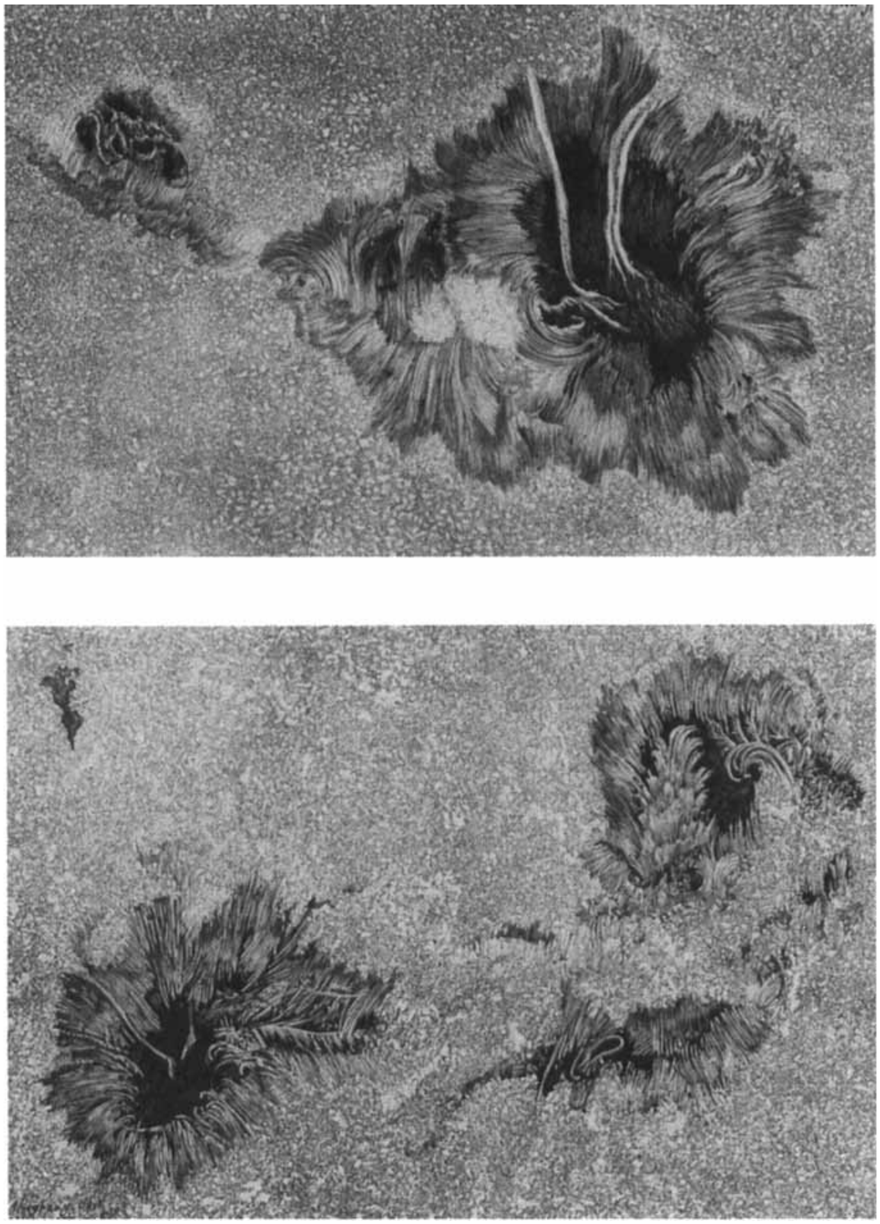

Fig. 1-3 Drawings of sunspots made by S. P. Langley on September 21, 1870 (top) and March 5, 1873 (bottom), using the 13-inch refractor at Allegheny Observatory. Note the light "bridge" across the umbra and the fine penumbral detail. Granulation can be seen in the photosphere outside the spots. From S. P. Langley, *The New Astronomy*, 1888.

The development of more compact and achromatic refractors in the 19th century eventually made possible visual observations of greatly increased resolution, such as the superb sunspot drawings by S. Langley shown in Fig. 1-3. But the records compiled by 17th-century observers continue even now to provide new insights into the sun's behavior. For instance, we find that spots became rare after about 1645 and remained so until around 1715. This broad minimum of sunspot occurrence was rediscovered by E. W. Maunder in the late 19th century and has since been named after him. Since the processes responsible for sunspot generation and for the solar differential rotation may be related, efforts have also been made to reconstruct the solar rotation rate and its latitude profile from the rotation of spots observed immediately before 1645 and after 1715, to check for a possible change during the Maunder Minimum. The data used included Hevelius' drawings, such as Fig. 1-2. More than 300 years after their discovery, the mechanisms responsible for the generation of sunspots, the relation of spots to faculae, and the differential rotation of the sun's surface continue to present problems at the forefront of solar research.

Toward the end of the 18th century, advances in stellar astronomy associated in large part with the work of William Herschel gradually led to the view that the sun is a star similar to the thousands of bright points seen in the night sky, only located much closer to the Earth. How much closer became clear in 1837 with the reliable measurement of the stellar parallax of 61 Cygni by F. Bessel. Study of the sun's physical structure and chemical composition as a valuable key to understanding the stars provided a powerful impetus to solar research in the 19th century that continues to the present day.

1.2 The Spectroscope and Photography

The spectroscope was the new instrument that enabled astronomers to carry the idea of the sun as a star to useful conclusions. After initial studies on the dispersion of solar light by Newton and others in the 17th and 18th centuries, J. Fraunhofer built the first spectroscope useful for quantitative analysis in 1814 and used it to measure the positions of over 500 of the remarkable dark lines seen in the solar spectrum. Fraunhofer's original spectrum is illustrated in Fig. 1-4. A certain amount of experimental work on the spectra of the sun, stars, and incandescent laboratory sources ensued. But it required the insight of G. Kirchhoff almost 50 years later to discover the simple laws governing the emission and absorption of light from solid and gaseous bodies.

The most important impact of Kirchhoff's work on solar research lay in his demonstration that the thousands of dark absorption lines observed in the spectrum of the solar disk have, in general, a one-to-one correspondence to bright emission lines in incandescent vapors observed spectroscopically in the laboratory. From these foundations, Kirchhoff and others showed that the sun's surface consisted of hot gases made up of elements found on Earth. This discovery surely ranks with the most far-reaching findings of natural science, since it gave valuable support to the

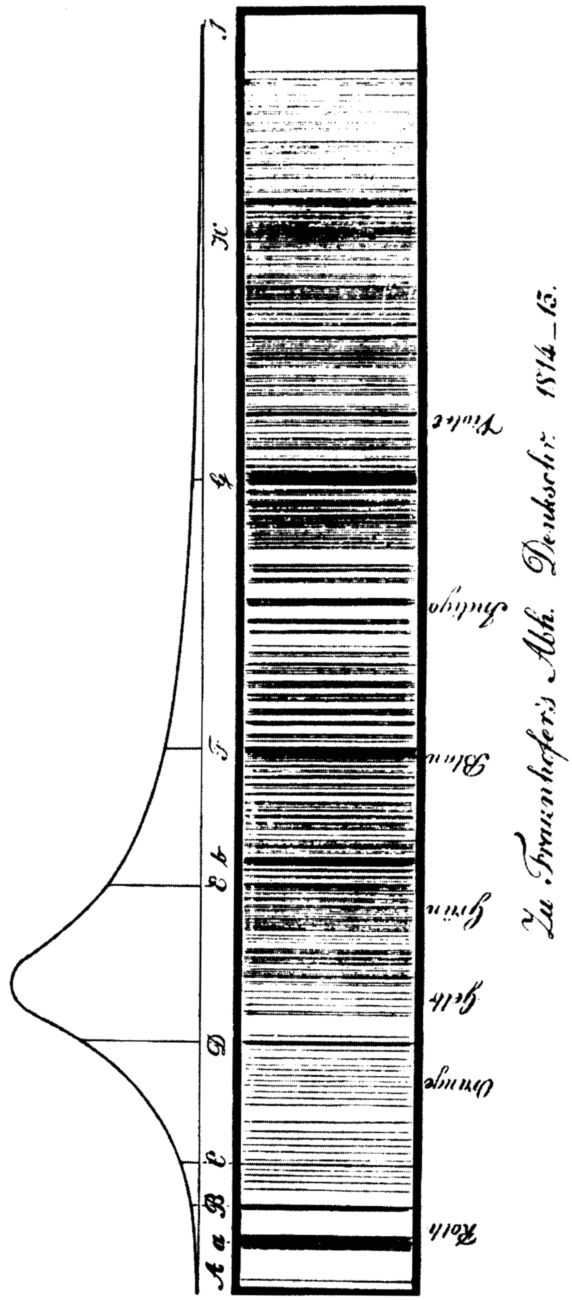

Fig. 1-4 Fraunhofer's original drawing of the solar spectrum from 1814. Red is to the left, violet to the right. The H-line of Ca II, the b-lines of Mg I, and D-line of Na I correspond to modern nomenclature. From J. N. Lockyer, *Solar Physics*, 1874.

rather daring cosmological principle that all corners of the universe obey the same laws of nature.

The development of photography in the 1840s made possible more objective recording of structure seen on the sun's disk in white light. One of the first results was to demonstrate that the darkening of the solar disk toward the limb claimed from visual observations was real. This limb darkening of the white-light-emitting layer in the sun's atmosphere (later called the photosphere) was eventually explained by K. Schwarzschild, in a classic paper in 1906. Schwarzschild used newly developed ideas of radiative transfer to show that the observed limb darkening by a factor of roughly 2.5 can be understood in terms of the decreasing temperature outward through the photosphere if energy transport through this layer is by radiation.

By the 1870s, photographs such as Fig. 1-5 were obtained of the pattern of small mottles, called granulation, that cover the photosphere. The bright roundish granules (also seen in the drawings of Fig. 1-3) are typically a few arc seconds in diameter and are separated by narrower dark intergranular lanes. It was later recognized that these granules resemble the pattern of convective cells found in H. Bénard's experiments of 1901 on fluids heated from below, so that they probably represent hot rising gas elements convecting heat to the sun's surface.

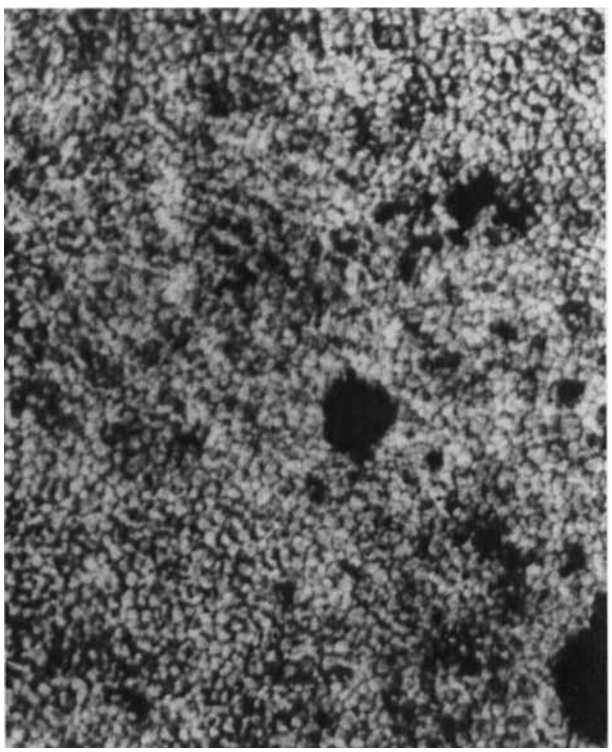

Fig. 1-5 A photograph of the photosphere, showing some spots and pores and perhaps granulation, taken in Paris by J. Janssen on July 5, 1885.

This association of granules with convection is generally accepted on the basis of modern observations and simulations, but the validity of the early evidence discussed above has been questioned. The remarkable regularity of the granulation seen in Fig. 1-5 seems to arise at least in part from a recently discovered pattern of fine cracks in the emulsion. The cells seen in Benard's experiments are considered to have been caused by surface tension effects rather than by convection. This ironic episode shows how compensating errors in observations and interpretation can underlie a correct result.

It appears that, although most of the energy is carried by radiation through the photosphere, the granulation seems to represent the "overshooting" of convective elements from the deeper, opaque layers of the sun where convection carries most of the heat flux. The structure of a deep convection zone, extending roughly 30 % of the sun's radius below the photosphere, was first worked out in 1930 by A. Unsöld.

1.3
Solar-Terrestial Research and the New Astronomy

A different line of inquiry that had great impact on 19th-century solar research, and continues to provide important motivation today, was initiated by the discovery of a significant 11-year periodic variation in the number of sunspots by H. Schwabe in 1843. Soon afterward, R. Wolf and E. Sabine connected these results to the modulation in geomagnetic storm occurrence discovered earlier by J. Lamont and by Sabine in research on magnetic disturbances in Germany and in Canada. This was the first good evidence for a variable solar influence on Earth, and the beginning of solar-terrestrial research.

More insight into the specific aspect of solar activity responsible for geomagnetic disturbances came after the first observations of great solar eruptions within spot groups, reported in white-light observations of the photosphere in 1859 by R. Carrington and R. Hodgson. The occurrence of a great geomagnetic storm within less than a day suggested that these "flares" on the sun, rather than the sunspots themselves, were directly responsible for the terrestrial disruptions. The impressive correlations found thereafter between geomagnetic storms and disk passage of large spot groups gradually led to our relatively recent acceptance of the view that emissions of charged particles travel from the sun to Earth.

Observations during total solar eclipses provided the only method to study the solar atmosphere outside the bright photospheric disk until the 1860s. The ability to predict eclipses dates back to Babylonian times, and improvements in technique introduced by the great astronomer Hipparchus of Rhodes in the second century B.C. increased accuracy to a few hours. The extended white-light corona had been noted in the early 18th century, along with naked-eye sightings of the pinkish coronal condensations that we now call prominences. Eclipse watchers of about that time had also remarked on the red light of the narrow layer called the chromosphere, visible briefly just as the moon covers and uncovers the photosphere at beginning and end of totality.

During the American Revolutionary War hostilities were suspended in a part of the state of Maine for one day to permit the Reverend Prof. Williams and his collaborators from Harvard College to observe the eclipse of 1780. This expedition was reportedly the first to note the phenomenon of Bailey's beads, which was rediscovered at the 1836 eclipse by the English amateur astronomer F. Bailey, and is caused by solar light shining through the lunar valleys at the edge of the moon's disk.

But until the event of 1836, few physical observations of the sun were made, although the times of eclipse contacts were systematically recorded. At eclipses beginning in 1842, some of the foremost astronomers of the day set up experiments intended to determine whether the corona, prominences, and chromosphere were solar or whether they originated in a (then – plausible) lunar atmosphere. Good photographs obtained at two sites 250 miles apart during the 1860 eclipse established that prominences were indeed some kind of cloud or condensation in the sun's atmosphere.

Kirchhoff's work on the interpretation of spectra encouraged spectroscopy at the 1868 eclipse. Here J. Janssen observed the yellow (D_3) line λ 5876 in prominences, later found to be radiated by neutral helium, an element not discovered on Earth until 1895 by the chemist W. Ramsay. The green coronal emission line at λ 5303 was also first observed during the eclipse of 1869. This was the first of 24 coronal emissions to be discovered in the visible spectrum which could not be identified with any known element, and for many decades posed a vexing problem for spectroscopy. The explanation came in 1941, when the spectroscopist B. Edlén, proceeding from earlier evidence put forward by W. Grotrian, showed how these lines ascribed to the mystery element "coronium" actually arose from forbidden transitions between low-lying fine structure levels of highly ionized heavy atoms such as Fe and Ca. In 1871, Janssen identified greatly broadened Fraunhofer absorption lines in the coronal continuum spectrum and showed that they originated in scattering of the photospheric Fraunhofer lines by coronal particles. He thus established that the white-light corona was solar in nature. Advances in understanding such as these, achieved through the application of spectroscopy to the sun and stars, came to be recognized as the New Astronomy, or astrophysics as we now call it.

The 1868 eclipse also led to the first use of monochromatic imaging of solar features, through the invention of the "spectrohelioscope" to observe prominences. Most of the emission from prominences occurs in a few narrow spectral lines, so the principle of this instrument was to use a prism spectroscope to form an image in a narrow spectral passband including one of these lines. Most of the continuum light of the solar disk scattered in the Earth's atmosphere was thus excluded. Using this technique, J. Janssen in France, and soon thereafter, J. Lockyer in England, were the first to observe prominences and also the chromosphere, outside of eclipse. Drawings of prominences and of the chromosphere made by W. Huggins and A. Secchi using spectrohelioscopes in the 1870s revealed structural details such as the chromospheric jets (seen at the limb in Fig. 1-6) that were rediscovered after 1945 and named spicules. These visual observers also documented the intricate shapes of prominences, such as the loops illustrated in Fig. 1-6, and their often rapid motions. Around this time, comparative study of the eclipsed corona near sunspot

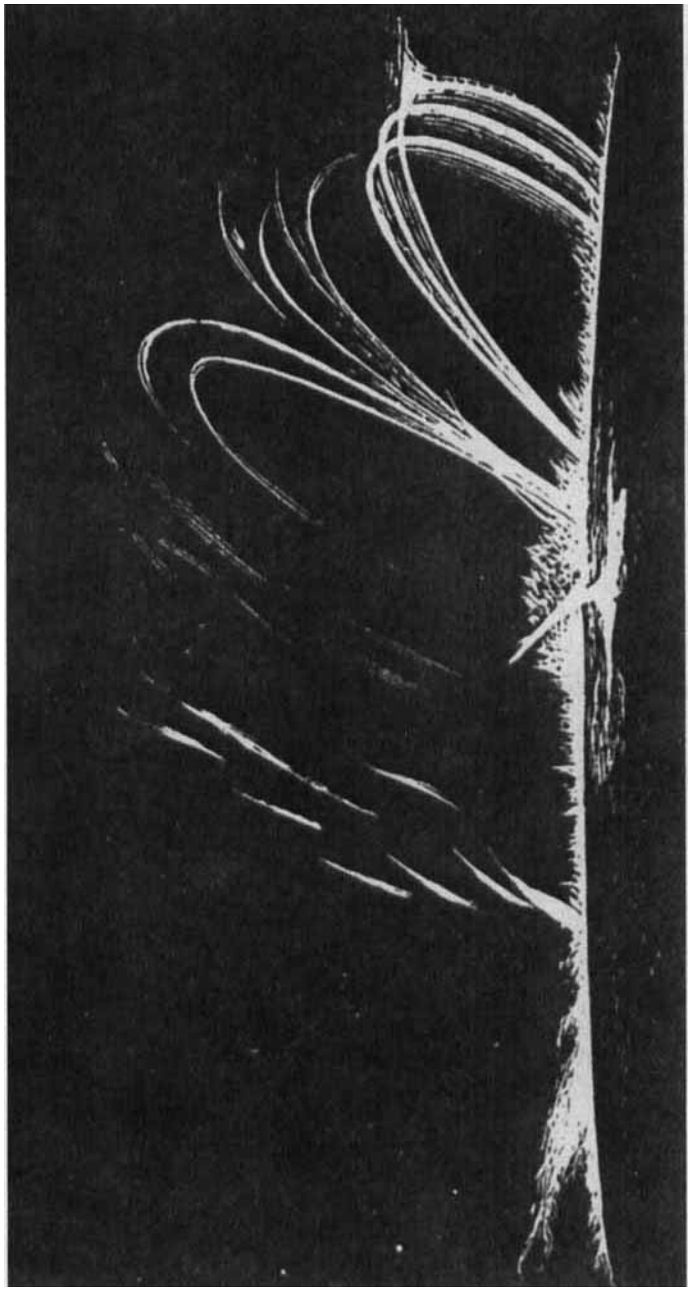

Fig. 1-6 Coronal loops drawn by A. Secchi from spectrohelioscope observations in the Balmer alpha line on October 5, 1871. The thin radial structures at the limb are chromospheric spicules. From C. Young, *The Sun*, 1895.

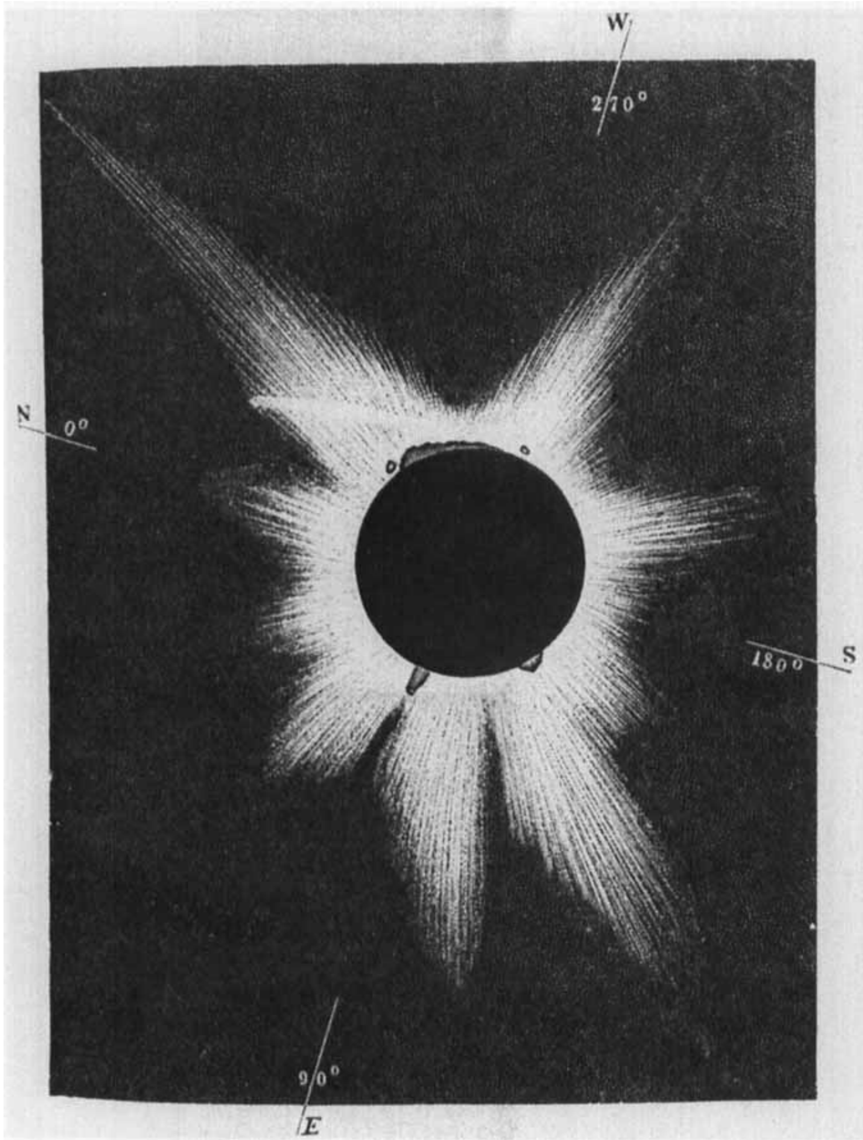

Fig. 1-7 A drawing of the corona of 1868 observed at Mantawa-lok-Kekee, Malaysia. This was the eclipse at which helium was first detected spectroscopically. Drawings of the same eclipse by different observers usually showed considerable variation. From J. N. Lockyer, *Solar Physics*, 1874.

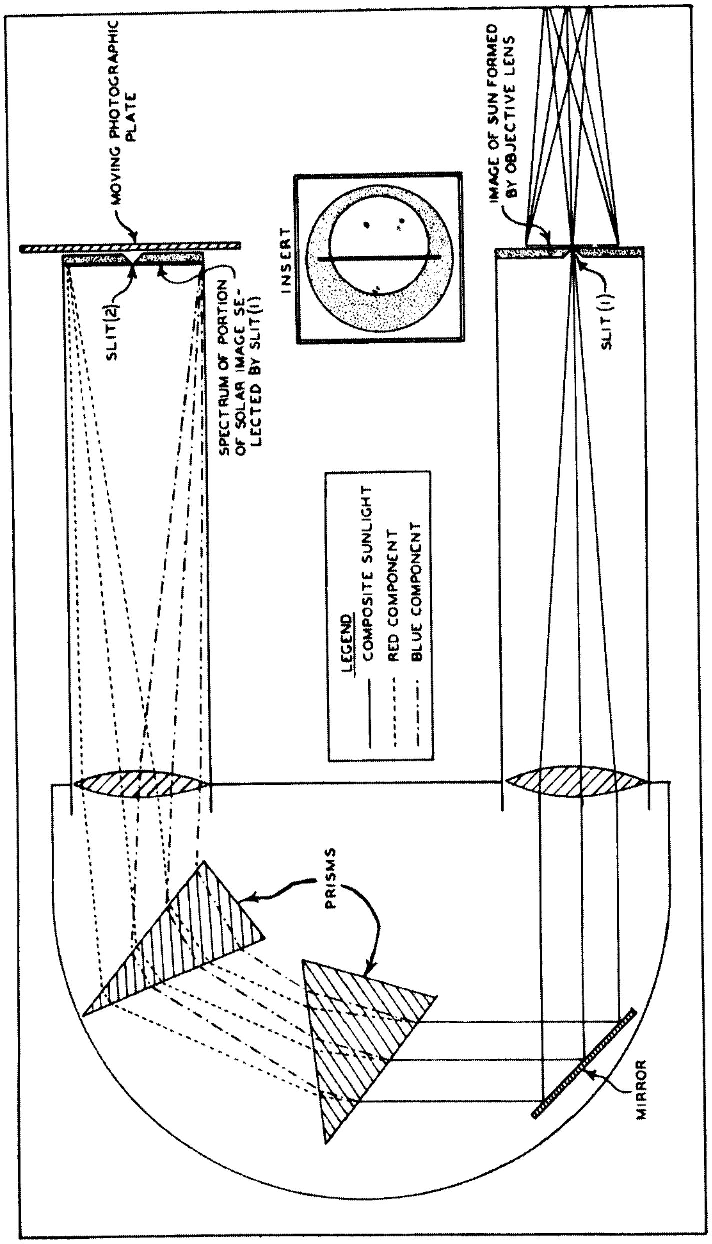

Fig. 1-8 The principle of the spectroheliograph. Slit 2 is adjusted in position and width to select a wavelength and pass-band width of interest. The photographic plate is then translated across slit 2 at exactly the same rate as the solar image is translated across slit 1. By permission of Mt. Wilson and Las Campanas Observatories, Carnegie Institution of Washington.

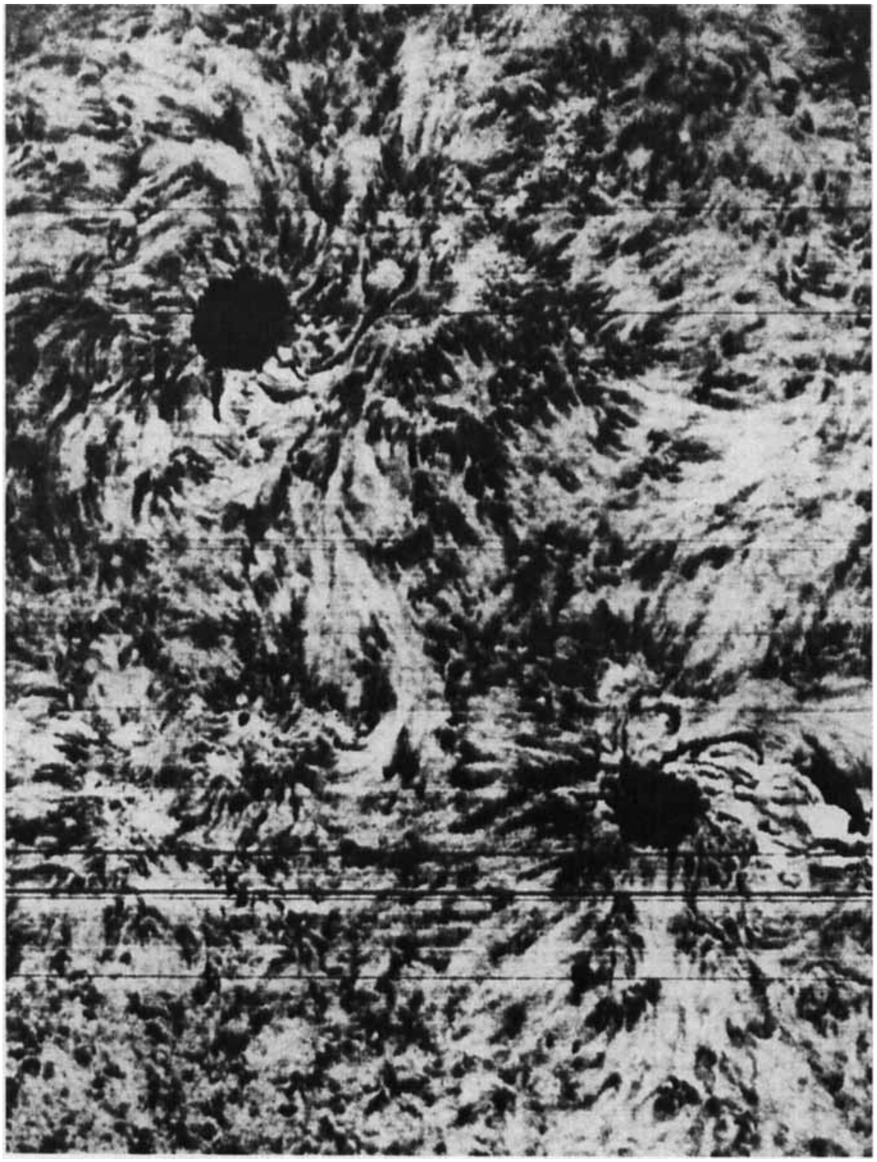

Fig. 1-9 The presence of intense magnetic fields in sunspots was suggested to Hale by vortex-like shapes in chromospheric structures similar to those shown in this Hα spectroheliogram taken September 9, 1908. By permission of Mt. Wilson and Las Campanas Observatories, Carnegie Institution of Washington.

cycle maximum and near the minimum (see Fig. 1-7) first revealed the differences in morphology characteristic of these periods of high and low activity. The intricate loops and plumes visible near maximum activity led to the first suspicion that electromagnetic forces play a role in the dynamics of the solar atmosphere.

At the eclipse of 1870, C. Young observed the "flash" spectrum of the chromosphere, which blazes out for a few seconds immediately before and after totality. In 1913, S. Mitchell showed that this spectrum could not originate in scattering of the Fraunhofer lines since it contains high excitation lines weak or absent in the photospheric spectrum. His reasoning also led to the first suggestion that the chromospheric temperature might exceed that of the photosphere. The identification in 1941 of the "coronium" lines with forbidden transitions in Fe X, Fe XIV, Ca XV, and other highly stripped ions proved that the temperature of coronal gases lying above the chromosphere was very high, over one million degrees. These high temperatures continue to stimulate research into the processes of nonthermal heating in the chromosphere and corona. Dissipation of waves or electric currents seems to be required since thermal heating of these layers by conduction, convection, or radiation from the cooler photosphere would violate the second law of thermodynamics.

The structure of the chromosphere on the disk was first revealed by observations with the spectroheliograph, developed independently in 1891 by H. Deslandres and by G. Hale. This instrument combined the narrow passband of the spectrohelioscope with the advantages of permanent and accurate recording made possible by the photographic plate. Its principle of operation (Fig. 1-8) relies on stepping the entrance slit of a spectrograph across the solar disk at exactly the same rate that the exit slit is moved across the photographic plate. A full image of the solar disk can then be produced in any spectral line or continuum region by setting the grating angle so that the spectrograph forms an image of the entrance slit at the exit slit in the desired wavelength.

This instrument made it possible to photograph the spatial structure of the chromosphere on the disk in the resonance line of singly ionized calcium at 3934 Å, known as the Ca II K-line. Later, observations were obtained in the strongest Balmer absorption line Hα at 6563 Å when red-sensitized emulsions became available in 1908. Hale's spectroheliograms in these lines, one of which is illustrated in Fig. 1-9, revealed the intricate dark structures of chromospheric fibrils and mottles, which later work related to the spicules that are seen in emission at the limb in Fig. 1-6. Hale's spectroheliograms also showed the bright chromospheric plages seen when faculae are observed on the disk in the cores of the Hα and Ca K-lines. Hale and Deslandres were nominated jointly for the Nobel prize for their invention of this remarkable instrument.

1.4 Solar Chemical Composition and Energy Generation

Kirchhoff made the first solar identifications of absorption lines of sodium, iron, magnesium, and other heavy elements around 1860. The presence of hydrogen in the sun's composition was revealed about ten years later by A. Angstrom and others. H. Rowland's development of excellent diffraction gratings led to his publication in 1897 of a superb atlas of the solar spectrum between atmospheric cutoff around λ 2975 and the visual limit near λ 7331. From these spectroscopic data, he was able to increase the number of elements identified in the sun to 36 by the end of the century.

The major advance to quantitative analysis of the solar line spectrum came in the 1920s after the development of N. Bohr's atomic theory and of M. Saha's ionization equation. This progress in atomic and statistical physics enabled H. Russell in 1928 to use Rowland's eye-estimates of line intensities to establish the rough relative abundances of elements in the solar atmosphere. The surprising result that the sun consisted mainly of hydrogen was by no means immediately accepted.

Around 1920, A. Eddington explained how hydrogen burning to helium might provide the energy to fuel the sun's luminosity. Building on earlier ideas, he suggested that the hydrogen fuel might be sufficient to account for the Earth's age implied by fossils. This age had been recognized by then to greatly exceed the $20 \cdot 10^6$ year time scale that H. van Helmholtz had estimated in 1854 for conversion of the sun's gravitational potential energy to radiation. Helmholtz's theory that a slow contraction of the sun provided the energy required to fuel its heat and light output remained the only reasonable explanation of the source of the sun's power output until about the turn of the century. The rate of steady contraction implied in the present epoch would be well below the detection limit for measurements of change in solar diameter.

Two sets of nuclear fusion reactions, called the carbon-nitrogen cycle and the proton-proton chain, were put forward in 1938 by H. Bethe, C. Critchfield, and C. von Weizsäcker to specify how the nuclear burning might proceed. The first direct test of these mechanisms in the sun was provided by the neutrino experiment of R. Davis, located deep in the Homestake gold mine in South Dakota. This experiment began operation in 1969 and detected a significantly lower neutrino flux than was predicted by standard solar models. The discrepancy is now known to result from then-unknown neutrino properties discussed in Chapter 6, and the solar models appear to be correct.

The success of nuclear fusion in explaining the luminosity of stars now rests largely on the good agreement between the observed and calculated distribution of stars in the Hertzsprung-Russell plot of stellar luminosity against color. Impressive regularities found in the relative solar abundances of elements heavier than helium are also well explained through the theory of stellar evolution put forward by M. Burbidge, G. Burbidge, W. Fowler, and F. Hoyle in 1957. Their explanation is based on successive fusion episodes, which build up the elements lighter than iron.

The sun's enormous power output was demonstrated by the first good calorimeter measurements of the total solar irradiance or "solar constant" obtained in 1837 by S. Pouillet in France and by J. Herschel working in South Africa. Pouillet's radiometer is shown in Fig. 1-10. Measurements of the spectral distribution of solar power were advanced by S. Langley's invention of the thermoconducting bolometer around 1881. With this sensitive detector, Langley mapped the sun's spectrum to 5.3 μm. He also showed the importance of molecules in the Earth's atmosphere, such as H_2O, CO_2, and CO, in the seasonally variable terrestrial absorption of solar infrared light.

C. Abbot, working with Langley at the Smithsonian Institution, put into use better radiometers for the solar constant measurements. Together with his collabora-

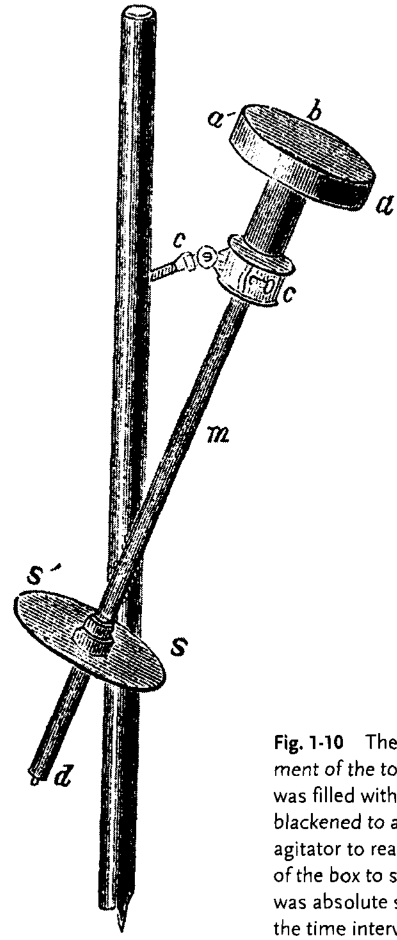

Fig. 1-10 The radiometer used by S. Pouillet in 1837 for measurement of the total solar irradiance. The hollow metal box marked *a-a'* was filled with a known volume of water. The upper surface *b* was blackened to absorb sunlight. The shaft *m* contained a thermometer-agitator to read the water temperature increase caused by exposure of the box to sunlight for a known interval of time. The measurement was absolute since the volume of the water, the blackened area, and the time interval were known.

tors, he then carried out an epic program of daily solar constant observations between 1923 and 1952 at a worldwide chain of Smithsonian mountain stations. Abbot's measurements demonstrated that the solar radiation transmitted by the Earth's atmosphere was constant to better than 1 % over several decades. But the absolute value of the solar constant remained uncertain at the 5 % level until the sun's ultraviolet flux below 3000 Å was properly measured by rockets in the 1950s.

Modern radiometry from space now provides the *absolute* value of the total solar irradiance to about 0.3% accuracy. Daily measurements since 1978 have clearly shown the decreases of the sun's output caused by dark sunspots, and the increases produced by bright faculae. These remarkably reproducible measurements, first achieved by R. Willson at JPL, and by J. Hickey at the Eppley Laboratories, also showed that the sun is almost 0.1% brighter, not darker, near peak sunspot activity. The explanation for this unexpected behavior is discussed in Chapter 13.

1.5
The Mt. Wilson Era of Large Telescopes

Using first the horizontal Snow telescope, and then the 60-ft and 150-ft tower telescopes built at Mt. Wilson near Los Angeles in 1907 and 1912 (Fig. 1-11), Hale and his collaborators made the first solar observations using high-dispersion spectrographs and large image scale to achieve good spectral and spatial resolution on the sun. Following the example of J. Lockyer, E. Frankland, and other chemists, they compared their solar spectra with those they obtained of laboratory plasmas at various temperatures and pressures. The laboratory at Mt. Wilson, where these comparisons were carried out, is shown in Fig. 1-12. The most important results had to do with the physics of sunspots and of the solar magnetic field. Hale's work confirmed (by comparison of the umbral spectrum with arc and spark spectra in the laboratory) that the dark umbra is cooler than the brighter photosphere. This fact was less obvious than it seems now, since the theory of radiative transfer in gaseous atmo-

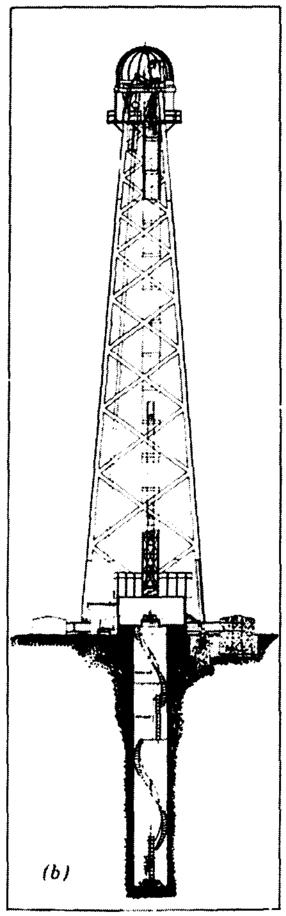

Fig. 1-11 (Left) The 150-ft tower telescope still in active use on Mt. Wilson and (right), cutaway view of the 150-ft solar tower. The deep pit below ground level contains the spectrograph, with its grating at the bottom. By permission of Mt. Wilson and Las Campanas Observatories, Carnegie Institution of Washington.

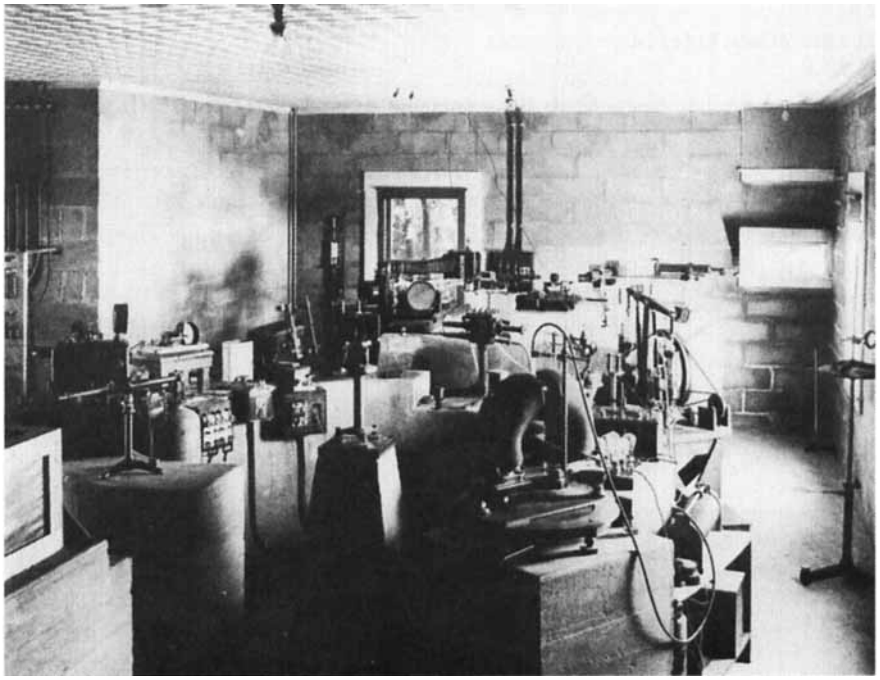

Fig. 1-12 Interior of early physical laboratory on Mt. Wilson, showing spectrograph and magnet used in studies of the Zeeman effect. By permission of Mt. Wilson and Las Campanas Observatories.

spheres was only beginning to emerge in the work of A. Schuster, K. Schwarzschild, and others.

High-dispersion sunspot spectra also showed a puzzling widening and sometimes splitting of the lines in umbrae. Hale's Hα spectroheliograms in 1908 had shown vortex-like chromospheric dark filaments wound around the spots (see Fig. 1-9) resembling the configuration of iron filings around a magnet, which led him to suspect strong umbral magnetic fields. Proceeding from P. Zeeman's (1896) analysis of the magnetic field splitting of spectral lines, Hale obtained umbral spectra through a Nicol prism and found the expected zigzag pattern caused by the opposite circular polarization of the Zeeman-split line-wing components in sunspot fields of up to 3,500 G.

This first discovery of extraterrestrial magnetic fields set the stage for the systematic study by Hale and S. Nicholson of the laws of sunspot polarity and their changing behavior over the spot cycle. Their work showed that the oscillation of sunspot number recognized since 1843 constituted half of a 22-year cycle of sunspot magnetic polarities. The most obvious aspect of this 22-year cycle is the reversal of polarity of the east and west spots in active regions that accompanies the onset of a new 11-year cycle in spot number.

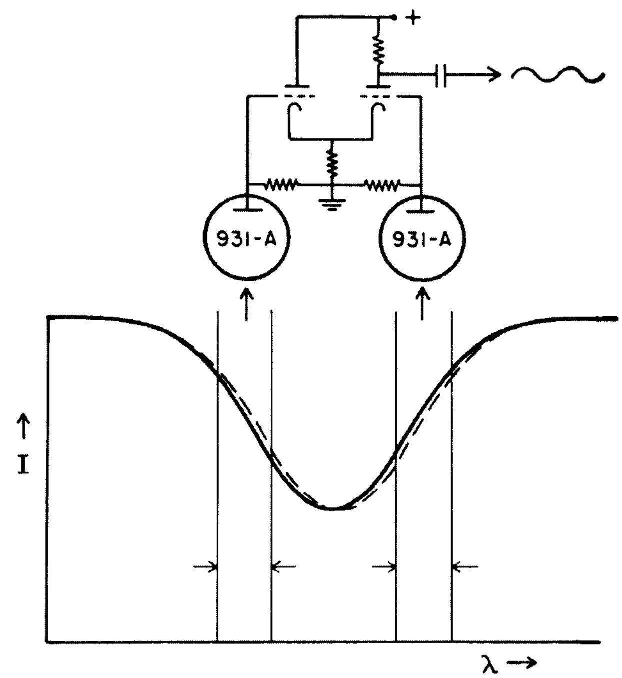

Fig. 1-13 Schematic diagram showing the principle of operation of the original Babcock magnetograph. Profiles of an absorption line for the two states of opposite circular polarization are shown solid and dashed. The symbols I and λ refer to intensity and wavelength respectively. The two detectors are type 931-A photomultipliers. By permission of the University of Chicago Press.

Hale's further evidence for a general polar magnetic field of intensity roughly 50 G with opposite sign at the north and south poles, was never fully accepted. Only after the development of the photoelectric magnetograph by H. W. Babcock in 1951 was a weak 1–5 G polar field reproducibly detected in spatial averages over the polar regions and observed to change in polarity every 11 years along with the spot fields.

The principle behind this very important instrument can be sketched as follows. Two exit slits of a high-dispersion spectrograph are placed in the red and blue wings of a magnetically sensitive photospheric absorption line, whose intensity profile is shown in Fig. 1-13. The solid and dashed $I(\lambda)$ curves denote the profiles from a magnetic region of the sun when left or right-hand circularly polarized light is alternately admitted to the spectrograph. Since the two wings of a normal Zeeman triplet broadened by a magnetic field are oppositely circularly polarized, admitting alternate senses of polarization by rotating a quarter-wave plate placed in series with a Nicol prism, moves the profile back and forth. The difference in intensity measured between the two photomultipliers placed in the wings changes from one profile to the other. This variation of the difference signal is amplified and recorded. Its ampli-

Fig. 1-14 An early photoelectric magnetogram obtained with the Babcock magnetograph. Each trace was made by a separate scan across the sun. A deflection equal to the distance between traces is produced by a field (averaged over the spectrograph entrance slit) of about 10 G. By permission of the University of Chicago Press.

tude is roughly proportional to the *net* magnetic flux (the effect of opposite magnetic polarities will tend to cancel) in the solar region whose light is admitted through the entrance slit of the spectrograph.

The magnetograph revealed for the first time, the distribution of photospheric magnetic fields outside of spots and faculae (Fig. 1-14). Observations with improved versions of this instrument by the 1960s showed a network pattern of magnetic fields of characteristic cell dimensions 20,000–40,000 km covering the quiet photosphere. This magnetic network was found to coincide well with the chromospheric network defined by the bright emission in Ca K spectroheliograms outside of active regions.

The suggestion of a velocity shift of spectral lines by C. Doppler in 1842 eventually led to the first measurements of the relative blue- and redshifts of Fraunhofer absorption lines at the sun's east and west limbs and thus to spectroscopic measurement of the photospheric gas rotation rate. These measurements (by N. Duner, H. Vogel, and C. Hastings between 1870 and 1890) showed a similar rotation rate and equatorial acceleration to the accurate sunspot rotation rates obtained by Carrington in 1865. But the plasma rates could be extended to latitudes ±75°, whereas spots could be used as tracers only over the active latitudes to roughly ±40°. Motions around sunspots also proved interesting. J. Evershed's 1909 observations in Kodaikanal, India, revealed a strong horizontal outflow extending several spot diameters from the umbral edge. This Evershed effect has yet to be incorporated into a satisfactory dynamical theory of spot coolness and stability.

The improved accuracy of line wavelength determinations achieved by upgrading the Rowland atlas in 1928 to use the cadmium red line interferometric wavelength standard led to increased awareness of the lirnb redshift of Fraunhofer lines relative to their disk center position. The nature of this shift, noted earlier by Hale and his collaborators at Mt. Wilson, generated lively debate. A small redshift relative to the laboratory wavelength, of constant magnitude across the disk, was expected from general relativity. But the conspicuous center-to-limb redshifts of up to 0.5 km s^{-1} required a different explanation. Only in the 1950s was a viable mechanism put forward in terms of the net Doppler blueshift near disk center caused by the bright upflowing granules, relative to the smaller contribution of dark downflowing material of intergranule lanes.

In the late 1950s, R. Leighton devised a technique for studying periodic velocity signals on the sun, using the spectroheliograph at the 60-ft Mt. Wilson tower. The slit was placed in the wing of a line, and the spectroheliograph was scanned across the sun and back taking P minutes in each direction. A negative of the scan in one direction was then overlaid in perfect spatial registration on a transparent positive of the scan in the opposite direction. Points on the sun that happened to be Doppler shifted in the same sense on both the negative and positive were registered strongly as dark or bright on the resultant print, whereas points whose Doppler shift changed sign during the time, cancelled to a neutral gray in the Doppler summation print. The variable time delay across the summed print was ideally suited to detecting oscillations of unknown period in the solar velocity field over the range of periods shorter than the maximum time delay $2P$.

Using this technique, Leighton and his graduate students, R. Noyes and G. Simon, detected and studied the 5-min oscillation of the photosphere. The fruitful interpretation of the oscillation as standing acoustic waves trapped in resonant cavities below the photosphere was made in the early 1970s, when the observations of F. Deubner and the calculations of R. Ulrich independently showed the rich mode structure of these oscillations when their oscillatory power is plotted in the plane of spatial wavenumber k versus temporal frequency ω. The position of the modes in the k–ω plane, their width, and their splitting yield valuable information on the structure of the sun's interior, on the dynamics of wave damping by turbulent solar convection, and on the sun's internal rotation profile.

The Doppler cancellation techniques developed by Leighton also showed that the network magnetic fields coincided roughly with the edges of an outflowing horizontal velocity field of roughly 0.5 km s^{-1} centered within each of the cells. These 20,000–40,000 km diameter velocity cells were called supergranular convection. The strong magnetic fields measured at their boundaries in magnetograms were thought to be intensified from weaker fields pushed to the edges by the ram pressure of the convergent supergranule flows. More recent magnetograms obtained with much higher spatial resolution have shown that the network magnetic fields consist of individual vertical flux tubes in which the magnetic field intensity is of order 10^3 G, thus similar to that of a spot umbra.

1.6
Advances in Coronal Physics and in the Theory of Solar Activity

Instruments used for coronal studies did not progress much for 50 years after the introduction of the spectrohelioscope at the 1868 eclipse. Only a few minutes per year were available to study the corona outside of the brightest prominences (which were visible outside of eclipse through a spectrohelioscope) using portable equipment set up at remote sites. The idea of blocking out the photospheric disk at the image plane of a telescope had occurred to many, but the difficulties of overcoming scattered light were not solved until B. Lyot implemented some ingenious precautions and built the first working coronagraph at the Pic du Midi, in France, in 1930.

The principle of the coronagraph is illustrated in Figure 1-15. The photospheric disk is imaged on a convex occulting disk at the prime focus of a refracting telescope and reflected out of the beam. One of Lyot's important advances was in placing a field lens (D) behind the occulting disk (C) to image the primary objective lens (B) onto a circular diaphragm (E). This diaphragm was made somewhat smaller than the objective lens image, so that rays originating at the edge (A) of the lens were blocked. Lyot had found that diffraction from the edge of the lens contributed a large fraction of the total scattered light.

Through this arrangement only coronal light gathered by the central part of the objective was used to form the final image of the corona. To reduce scattering in the lens glass itself to a minimum, Lyot used a single lens of bubble-free glass instead of an achromatic doublet, which causes multiple internal reflections from the glass-air surfaces (unless modern antireflection coatings unavailable in 1930 are used). Finally he is reported to have taken care to clean the objective with oil from the tip of his nose, applied with a well-laundered handkerchief. A large coronagraph is shown in Fig. 1-16.

Lyot also co-invented the birefringent filter (with Y. Öhman of Sweden), which enabled him to take monochromatic pictures of coronal and chromospheric phenomena far more rapidly than could be done with a spectroheliograph. The principle of the filter relies on the interference produced between the ordinary and extraordinary rays passing through a birefringent crystal such as calcite. Since the refractive index of the crystal differs for the two rays, they emerge with a phase difference

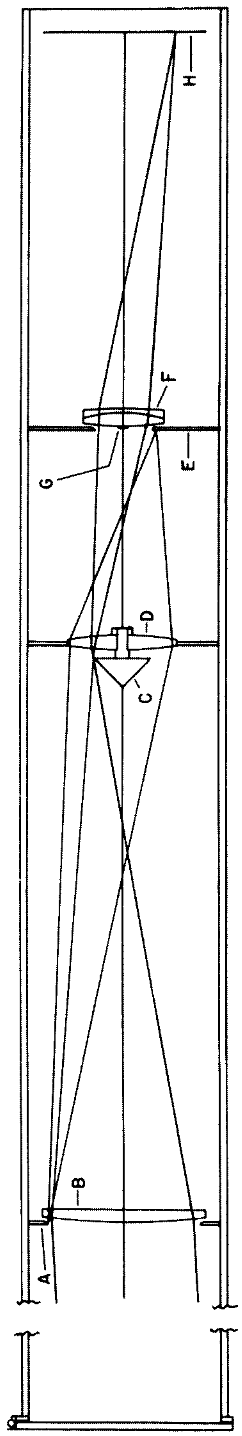

Fig. 1-15 Optical parts of a Lyot coronagraph. See text for explanation. By permission of the University of Chicago Press.

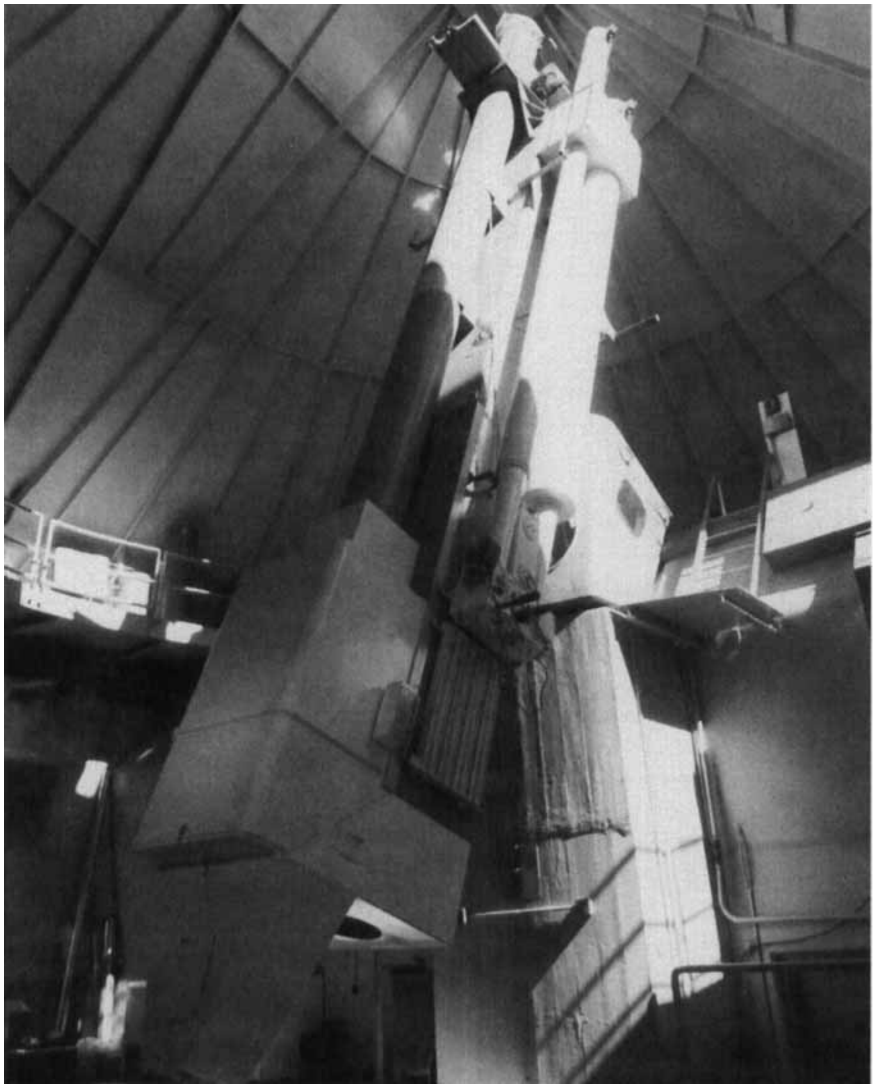

Fig. 1-16 The 16-inch aperture coronagraph at Sacramento Peak Observatory in New Mexico. This instrument is in active use for coronal studies and other observations requiring low scattered light. National Solar Observatories photograph.

that depends (for a given calcite thickness) on the wavelength of the light. Transmission is minimal for wavelengths at which the block of calcite yields a phase difference of π. A practical filter consists of several calcite elements, the thinner elements remove unwanted transmission sidebands at wavelengths other than that selected. The birefringent filter can be built to produce narrow passbands of 0.25 Å or less,