



Holly Gilbert is the director of the High Altitude Observatory at the National Center for Atmospheric Research in Boulder, Colorado. She previously served as the director of the heliophysics science division at NASA's Goddard Space Flight



Holly Gilbert

Even as our understanding of the Sun has grown, many fundamental questions remainsome of which have big implications for life on Earth.

ur Sun has been an object of mystery and study for as long as humans have been around to squint at it in wonder. Sunspot observations have been recorded for over 400 years, and eclipses were first documented thousands of years ago. Although the Sun is no longer a complete mystery, the way we observe it continues to develop, as continued technological advances uncover new mysteries for future generations to solve.

ADVANCES IN SOLAR TELESCOPES



FIGURE 1. LARGE ARCS OF SOLAR MATERIAL, or prominences, are common features of an active sun. They were only discovered once astronomers viewed the Sun through telescopes. The *Solar Dynamics Observatory* now provides high-resolution monitoring of such solar activity. (Courtesy of NASA's Goddard Space Flight Center Scientific Visualization Studio and the *Solar Dynamics Observatory*.)

The invention of telescopes facilitated more detailed study of sunspots and the identification of exotic structures suspended above the solar surface (see figure 1). More recently, the space age brought about a new era of solar observations, leading to groundbreaking discoveries about solar magnetism and its reach. Launched in the 1970s, *Skylab* was the first space station, and although not designated to be a solar mission, it opened our eyes to large eruptions originating in the outer layer of the solar atmosphere, called the corona. Since the launch of the *Solar and Heliospheric Observatory (SOHO)* in the 1990s, a joint project between the European Space Agency and NASA, the trajectory for solar physics discoveries has swiftly escalated. Despite advances, many fundamental questions remain regarding all regions of the Sun, including the solar interior and the storms released in the outer solar atmosphere.

Pesky problems in solar physics

The Sun experiences an 11-year activity cycle. During its peak, called the solar maximum, the number of sunspots and related eruptions increases significantly, at times averaging three storms a day. In the solar-minimum phase, there can be no sunspots for long periods of time. The Maunder Minimum was a seven-decade period in the 17th and 18th centuries during which sunspots became extremely rare. From 1672 to 1699, observations revealed fewer than 50 sunspots, compared with the thousands of sunspots typically seen over similar time spans. Astronomers still don't know what causes the 11-year solar cycle, let alone what led to the unusual lack of sunspots during the Maunder Minimum.

George Ellery Hale (1868–1938) and collaborators determined that sunspots and their 11-year cycle were of a magnetic nature in the beginning of the 20th century. Although we now know plasma flows and magnetism are at the core of that phenomena, the solar dynamo driving those fluctuations is still being studied and modeled. A missing piece of information that would help advance the modeling efforts is the Sun's global magnetic field, especially data covering the polar regions. To date, the polar regions of the Sun have never been observed since the poles are not easily seen from Earth.

Solar flares are also not fully understood. Among the most

explosive phenomena in the solar system, they were discovered in 1859 during the landmark Carrington event.¹ During his regular monitoring of sunspots, amateur astronomer Richard Carrington observed rapidly brightening patches of light near a sunspot group; the changes in brightness were visible without any advanced technology. That flare was associated with strong effects seen on Earth shortly after: Telegraphs caught fire, and auroras were seen as far south as the Caribbean. It proved that solar events could impact Earth. Being able to predict solar flares of that scale is more important now than ever before because of our increased reliance on technology.

Coronal mass ejections (CMEs) are another solar phenomenon powered by magnetic fields. Associated with flares and driven by magnetic reconfigurations, they were first observed from the ground in 1971, and their dynamics were elucidated by *Skylab* observations a few years later. Magnetic field interactions, especially in the corona, are at the root of eruptions

and flares, but exactly how and where fields interact is still being debated. There are several models that connect magnetic topologies to observations of erupting materials and fields. The coronal magnetic field is also the energy source and organizing principle behind solar-wind acceleration² and coronal heating, two related and complex physical phenomena.

The coronal-heating problem is perhaps the longest standing, most frustrating issue yet to be resolved in the solarphysics community.3 In daily life, you expect that you will feel cooler as you move farther away from a heat source. But the Sun's corona—despite being the outermost layer of the solar atmosphere—is a million kelvin hotter than the Sun's visible surface, the photosphere (see figure 2). Moreover, that temperature gradient occurs across a very short distance of about 100 km, called the transition region. Recent satellite missions (see the article by Leonardo Di G. Sigalotti and Fidel Cruz, Physics Today, April 2023, page 34) have partly solved the mystery of why the outer layer is the hottest. Once energy from the core reaches the surface, it is transported through the chromosphere, transition region, and corona via magnetohydrodynamic waves or magnetic-field-line braiding and subsequent reconnection.

Magnetic reconnection occurs when oppositely directed magnetic field lines in a plasma break and rearrange. That process releases a large amount of energy that can heat the plasma and accelerate energetic particles. Magnetohydrodynamic waves, the other method of energy transport, access energy in the subsurface convection zone and transport it through plasmas. In the corona, to heat the plasma that wave energy must dissipate a nontrivial amount of heat because of the behavior of damping time scales in plasmas. It is widely accepted that dynamic magnetic fields play a fundamental role in those processes, but astronomers don't yet have enough information to accurately model the complex nature and relative role they play in heating the solar atmosphere.

Obtaining measurements of the coronal magnetic field vector and inferring properties of the fields are essential steps in understanding the physics of such phenomena. Solving those outstanding problems—or at least significantly advancing our understanding of them—is of particular importance in the con-

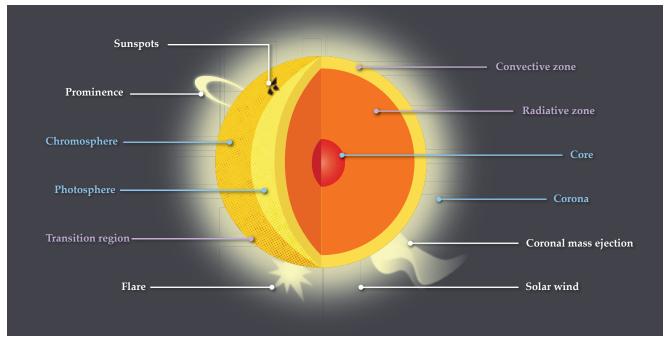


FIGURE 2. THE OUTER LAYERS OF THE SUN are the easiest to study, but even those still hide mysteries. The causes of surface phenomena (in white) such as flares and coronal mass ejections lie in the interior layers, which are harder to access. (European Space Agency image adapted by Abigail Malate.)

text of space weather. Our strong technology dependence in the modern era makes us increasingly vulnerable to the effects of Sun–Earth interactions during solar storms, especially energetic charged particles and magnetic fields carried great distances in CMEs. Those interactions affect the power grid, communications and other satellites, airlines flying over the poles, navigation, and astronaut health, making space weather a subject of national priority. As recently as 2020, PROSWIFT—a US bill to promote research to improve space weather forecasting—was signed into law.

Observational needs

A basic tenet of observational science is the constant need for improved spatial (and in the case of solar physics, temporal) resolution. You'd be hard-pressed to find a scientist who wouldn't advocate for it, especially in the solar-physics community. The physics and dynamics of the Sun necessitate a multiscale problem that can only successfully be understood with improvements to observations of solar activity. With each new mission or telescope, improved spatial and temporal resolution leads to new discoveries. Often the field of view (FOV) must be compromised to accommodate the extra resolution capability. Therefore, of equal importance is the existence of comprehensive, full-Sun images to provide the large contextual view. Large FOV data sets capture the global consequences of processes that occur on short time scales and small spatial scales.

Because the Sun is a three-dimensional sphere, multiple view-points are essential for obtaining accurate information about its inherent magnetic and thermal nature. Helioseismology (see the article by John Harvey, Physics Today, October 1995, page 32) allows a unique view into the solar interior using methods similar to seismology on Earth. A few space missions, such as the *Solar Terrestrial Relations Observatory*, have flown in

special orbits designed to collect data from 360° around the Sun. The *Parker Solar Probe* (see the article by Nour E. Raouafi, Physics Today, November 2022, page 28) and *Solar Orbiter* fly extremely close to the Sun—within the solar corona. Other mission concepts are being developed to cover still more regions of the Sun. Together, those types of extended viewpoints provide a global context that is critical to understanding the basic physical processes driving solar activity.

A common theme behind the unsolved solar-physics problems is the need for magnetic field measurements. But those measurements remain difficult to make, and the inferences drawn from existing observations are most mature for the photosphere. In the higher layers of the atmosphere, however, advances continue to be made in inferring the magnetic field through spectropolarimetry, the measurement of polarization as a function of frequency.

Detection methods

Studying light from the Sun involves more than just counting how many photons are detected and in which filters. Coronagraphs, which occult the bright disk of the Sun, allow observations of the diffuse corona. Polarimeters, which measure the polarization of light, can be used to infer information about the magnetic field and composition of the Sun. Both instruments select components of the total light from the Sun to gain a better understanding of its dynamics.

White-light coronagraphs allow glimpses of the coronal plasma and open observational windows to some of the most important causes of space weather: CMEs. Invented by astronomer Bernard Lyot in the 1930s, the coronagraph was innovative due to its aperture—called the Lyot stop—which blocks light diffracted around the entrance aperture. Coronagraphs work by allowing light to enter the telescope aperture as an

ADVANCES IN SOLAR TELESCOPES

evenly illuminated source. They require the rejection or suppression of stray light to a very high degree using a properly designed occulter system made up of baffles, stops, an internal occulter, optical coatings, out-of-band light rejection, and a highly polished objective lens.

Figure 3 shows a coronagraph design in which the secondary mirror at the center blocks the light with a lens that subsequently images it. Where a camera or detector would usually record the image, an occulting spot, also called a focal plane mask, is placed instead. That absorbs most of the light from the center of the FOV, while the telescope pupil is reimaged by another lens. The remaining light from the central source is concentrated around the edges of the pupil, forming rings around the edge of the aperture image and the secondary mirror image. The goal is to block as much unwanted light as possible by the time it reaches the detector. The next-generation coronagraphs have the capability to use spectropolarimetry to study flows, waves, density, and magnetic fields.

Polarization of light is expected whenever the mechanisms producing the observed radiation—whether radiative or collisional act in the presence of symmetry-breaking conditions. For example, the scattering of Planck-

ian blackbody radiation in a homogenous and isotropic gas will be unpolarized. The same process, however, will produce polarized light if the gas is illuminated nonisotropically, where the bulk of radiation comes from a preferential direction. Polarized light is seen, among other places, in the scattering of solardisk light by coronal structures observed off the solar edge.

Another typical symmetry-breaking situation producing polarized light is the radiative emission by atoms in the presence of an ambient magnetic field. In those cases, our ability to model the production of polarized light and interpret its signature in our data allows us to diagnose the thermodynamic and magnetic conditions of the emitting gas. Solar scientists largely rely on theoretical, numerical, and experimental polarimetric tools to diagnose the magnetism of the solar atmosphere.

The characterization of polarized light requires the specification of additional radiation quantities beyond intensity. Two parameters are needed to fully specify the degree and direction of linear polarization. A third parameter specifies the state of circular polarization about the propagation direction. Light detectors, however, are typically only sensitive to intensity signals. Hence, a polarimeter generally processes the incoming polarized light through some variable, birefringent optic system (the modulator) that encodes the polarization information into a finite series of varying intensity signals that can be captured with our detectors after being filtered by a linear polarizer (the analyzer).

Observing from the ground

Despite the availability of satellites, ground-based solar observatories continue to be an important source of solar imagery, quantitative spectroscopy, and polarimetry. Remaining on the

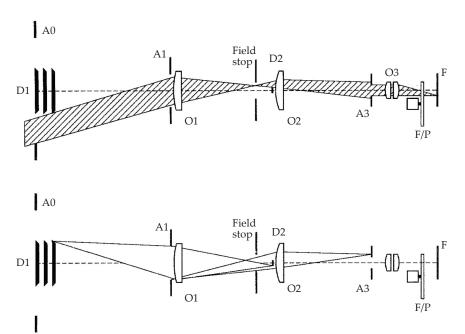


FIGURE 3. THE LARGE ANGLE AND SPECTROMETRIC CORONAGRAPH C3 on the Solar and Heliospheric Observatory (SOHO) satellite. The top diagram illustrates image formation; the bottom diagram shows how stray light is blocked from reaching the image plane. Sunlight enters from the left. The components shown are the front aperture (A0), external occulter (D1), entrance aperture (A1), objective lens (O1), field stop, internal occulter (D2), field lens (O2), Lyot stop (A3), relay lens with Lyot stop (O3), filter and polarizer wheels and shutter (F/P), and focal plane (F). (Adapted from ref. 10.)

ground offers multiple advantages, one of the most obvious being cost efficiency. Going to space is expensive and limits the ability to service and maintain instruments, whereas on the ground, it is easy to physically access the instruments for repairs and upgrades. Moreover, there are limits to the size of the telescope aperture and the complexity of the instruments that can be afforded in space. The largest solar telescope in space is the 50 cm *Hinode*, which is an order of magnitude smaller than the newly functioning ground-based Daniel K. Inouye Solar Telescope (DKIST), discussed below.⁴ DKIST offers unprecedented views of the Sun that would be impossible to obtain in space because of cost limitations.

Ground-based white-light coronal instruments have been operating at the High Altitude Observatory site in Climax, Colorado, from 1956 to 1963 and at NSF's Mauna Loa Solar Observatory (MLSO), located on the island of Hawaii, since 1965. Although limited by weather and the nighttime hours, a series of coronameters and coronagraphs (each instrument an improvement over the previous one) has provided important information on the innermost part of the corona and on the origins of CMEs. According to models, those large eruptions are often initiated by magnetic reconnection in the low corona, so observing close to the solar surface is advantageous, something that the design of the MLSO white-light coronagraph enables (see figure 4). The Coronal Multi-channel Polarimeter (CoMP) located at MLSO and its upgraded follow-on, UCoMP, are another type of ground-based coronagraph. UCoMP offers a unique combination of magnetic diagnostics, thermal information, and Doppler motions because of its sampling of multiple coronal lines, and its new filter is designed for a very broad bandpass that covers a range of 5000 Å.

UCoMP is a prototype for a larger telescope being developed as part of the Coronal Solar Magnetism Observatory, a proposed suite of complementary ground-based instruments (funding and location to be determined) designed to study magnetic fields and plasma conditions in the Sun's atmosphere. The primary instrument is a large coronagraph (the next-generation UCoMP) that offers better resolution, polarization measurements, and the ability to obtain line-of-sight magnetic fields (UCoMP only provides plane-of-sky magnetic fields). The suite's supporting instrumentation—a chromosphere magnetometer and a white-light coronagraph—measure magnetic fields in the Sun's chromosphere and the density of electrons in the corona. Together, those provide new tools such as vector-field measurements and invaluable predictive clues about damaging solar events.

The larger a telescope is, the more photons it can collect and the higher resolution it can achieve. NSF's DKIST on the Haleakalā volcano on Maui (see page 40) is now the largest solar telescope in the world. It was designed as a coronagraph that enables coronal spectroscopy and magnetometry. It collects more sunlight than any solar telescope and provides the best resolved and sharpest images of the Sun. DKIST's innovative technology includes an off-axis optical system with a large (4 m) primary mirror, active and adaptive optics, advanced optical and IR instruments, and versatile light-distribution optics that facilitate the simultaneous use of multiple instruments.4 DKIST's instrument capabilities are centered on spectropolarimetry. The ability to precisely measure the magnetic field throughout the solar atmosphere, including the corona, enables DKIST to address basic research aspects of space weather. DKIST will operate for two solar cycles and conduct communityproposal-driven observations.

An even more powerful tool being planned is a network of ground-based instruments at various longitudes designed to obtain wide coverage of observations. NSF's Global Oscillation Network Group (GONG) has served a vital role in comprehensive measurements of solar oscillations and magnetic field measurements in the photosphere since it was built in the mid 1990s. The six-site network (Australia, India, Canary Islands, Chile, California, Hawaii) was not designed with space weather in mind; that requires better sensitivity and new capabilities. The next-generation GONG is being developed to fill that critical gap once GONG's lifetime expires sometime in the 2030s, owing to aging instrumentation and the difficulty of replacing old hardware that no longer exists for purchase. Through a partnership of NSF institutions (the National Solar Observatory and National Center for Atmospheric Research) and potentially other government agencies, the next-generation GONG will install at each site some combination of spectropolarimeters for the precise measurements of solar magnetic fields at multiple heights, coronagraphs capable of monitoring the violent ejecta of magnetized plasma from the Sun's atmosphere and determining coronal magnetic topologies and plasma properties, and instruments for Doppler-velocity measurements required for helioseismology studies.

Another potential ground-based network of solar instrumentation is being explored by a group of international observatories. They are currently carrying out a preliminary design study of a synoptic solar-observing facility called the Solar Physics Research Integrated Network Group, funded by the

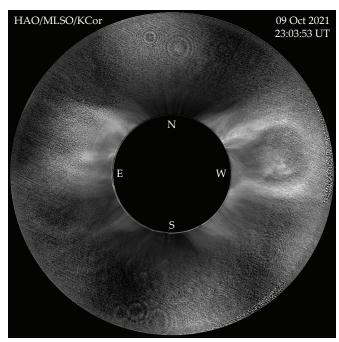


FIGURE 4. A WHITE-LIGHT IMAGE from Mauna Loa Solar Observatory's K-coronagraph shows a coronal mass ejection on the right side of the masked solar disk. The eruption occurred on 9 October 2021. A video of the coronal mass ejection is available at the online version of this article. (Courtesy of HAO/MLSO.)

European Union.⁶ The data products will include images with arcsecond resolution in multiple wavelengths, synoptic data including vector magnetic fields and surface velocity, and observations of flares and other transient events.

Other methods based on radio emissions provide constraints on the magnetic field in the corona by utilizing multiwavelength observations that are sensitive to both thermal plasma and nonthermal particles in addition to solar magnetic fields. The science of solar radio astronomy is many decades old. The Owens Valley Solar Array, established in 1979 and located near Big Pine, California, performs radio interferometry at many frequencies and conducts microwave imaging spectroscopy. It can detect thermal radiation in the chromosphere and corona and nonthermal radiation from solar flares. The Frequency Agile Solar Radiotelescope is a next-generation radio telescope that is being proposed to NSF (location and funding to be determined) and is specifically for solar observations.^{7,8} It's designed to perform Fourier synthesis imaging, exploiting the Fourier-transform relationship between the quantity measured by an interferometer and the radiobrightness distribution, and will combine ultrawide frequency coverage, high spectral and time resolution, and excellent image quality. That approach to solar observing offers a unique capability of measuring coronal fields against the solar disk.

Observing from space

The next generation of ground-based telescopes provides spectacular opportunities for new solar data and a chance at making exciting discoveries, but in the present age, ground-based data are made better when combined with space-based data. The obvious downside to observing from the ground is the

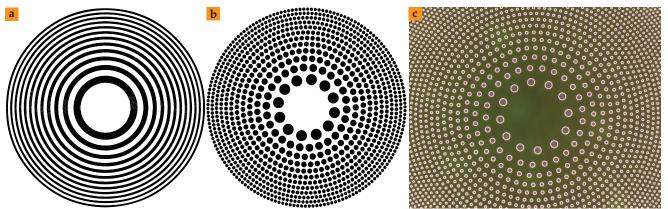


FIGURE 5. ZONE PLATES are one way to image the Sun near the diffraction limit. (a) A classical Fresnel zone plate has concentric rings with decreasing widths. (b) The photon-sieve variant has concentric rings of circles with decreasing radii. (c) A fabricated photon sieve needs to be precise to achieve the high-resolution images desired. The one seen here has holes as small as 2 μm in diameter. (Panels a and b courtesy of Adrian Daw; panel c courtesy of Kevin Denis.)

inability to observe more than what weather and daytime permit. Also, an important part of the electromagnetic spectrum—including the extreme ultraviolet (EUV) and x-ray regions, both important to seeing the dynamic and explosive activity on the Sun—is not accessible through Earth's atmosphere.

The space age brought with it groundbreaking discoveries in solar physics. One such discovery was *Skylab*'s sequences of images of CME dynamics in the 1970s. Multiwavelength observations, including those only accessible from space, opened windows into the various layers of the solar atmosphere. The limitations of space-based observations include cost and challenging access for equipment repairs. Yet the pros far outweigh the cons, and we have become incredibly adept at building resilient spacecraft and instrumentation, partly due to rigorous processes involved in delivering a proof of concept. Prior to being commissioned for spaceflight, many solar instruments are flown on suborbital sounding rockets to stress test them or on balloons. The long process of getting instrumentation to an appropriate level of technical readiness for space somewhat guarantees low risk and high success rates.

Multiple solar coronagraphs have flown on suborbital platforms since the 1960s, and the advances within the last couple of decades have significantly contributed to better forecasting capabilities for space weather. Because of the stringent size and weight limitations of spaceflight, designers of spaceborne coronagraphs face difficult challenges when trying to eliminate stray light. Those instruments employ multiple external occulters with shapes optimized to minimize the effects of scattered light. For researchers to view the low corona, the external occulters must be located at a very large distance from the objective lens. That makes it challenging for space-based externally occulted coronagraphs to match the FOV of the groundbased K-coronagraph at MLSO, about 1.05 to 3 solar radii.

Space-based solar coronagraphs that use multiple occulters, however, can get a better signal-to-noise ratio farther away from the solar disk and are capable of extending the outer FOV to large distances away from the solar surface; *SOHO*'s Large Angle and Spectrometric Coronagraph C3, for example, extends out to 33 solar radii. Moreover, the almost-continuous observing facilitates the forecasting of Earth-directed CMEs and other activity. The newest conceptual coronagraphs de-

signed for space explore a range of the spectrum not accessible to the ground, offering the potential of obtaining magnetic field measurements through spectropolarimetry. Some newer concepts separate the occulter from the telescope, introducing unique challenges with stability and jitter. Proba-3, a European Space Agency mission planned to launch in 2024, will feature two spacecraft—one an occulter and one a coronagraph—flying in formation 144 m apart, with millimeter accuracy, to create an externally occulted Lyot-style white-light solar coronagraph.

The improved technology, and subsequent rise in popularity, of CubeSats has provided some flexibility in flying instrumentation. CubeSats are small satellites that use a standard size and form factor. The field of heliophysics has been increasingly capitalizing on the more cost-efficient access to space they offer. CubeSats, however, require small instruments, so there has also been a push toward the miniaturization of hardware. Some observing capabilities currently cannot be shrunk beyond a certain threshold, but innovative coronagraphs have decreased in size to be flown with other instruments. Small coronagraphs can utilize folding mirrors or a boom to deploy the occulter and thereby reduce the size of the instrument volume on the spacecraft. That is particularly useful when sending instruments to nontypical orbits in space. Having observations from other viewpoints off the Sun-Earth line of sight is incredibly useful for seeing CMEs directed at Earth and obtaining more accurate information about their speed and location.

Observing only from the Sun–Earth line forces us to look at the Sun as a 2D object, and we miss out on some important activity on its backside and have incomplete information on the magnetic field, which leads to inaccuracies in models of the corona and solar wind. A 360°, synoptic view is ideal. A constellation of small satellites deployed around the Sun and over its poles could get us closer to that ideal. This would also help us understand fundamental and universal processes that apply to other stars.

An unconventional look forward

Considerable indirect evidence suggests that the milliondegree solar corona is heated by releases of energy on scales smaller than 100 km, which is about 0.15 arcseconds as seen from Earth. High-temperature solar plasma emits radiation primarily at EUV and x-ray wavelengths, which must be observed from space. But no conventional reflecting telescope has been fabricated with the extreme accuracy and smoothness necessary to achieve diffraction-limited imaging at those wavelengths. EUV wavelengths would require subnanometer accuracy, which is extremely difficult to achieve for meter-scale mirrors. A flat diffractive optic, such as a Fresnel zone plate (see figure 5), has been demonstrated to produce nearly diffraction-limited images with a tolerance at least an order of magnitude looser than conventional optics.

A conventional optical surface such as a mirror or lens is a powerful way to bend light rays. But each surface must have just the right tilt with respect to each incoming ray to produce a focal point. A zone plate can be pictured as transforming each incoming ray not into a single deflected ray but instead into a cone of outgoing rays, only some of which interfere constructively to produce a focal point. An EUV space observatory with a diffractive optic of modest aperture (less than 1 m in diameter) can probe the small spatial scales at which coronal heating is believed to take place. Recent progress suggests a suitable diffractive optic should be available in the near future.

A variant of a zone plate is the photon sieve. A photon sieve 80 mm in diameter has been shown in the lab to produce nearly diffraction-limited EUV images. A mission concept in development calls for 170-mm-diameter sieves to achieve the angular resolution required for its scientific goals. The most challenging aspect of that coronal microscale observatory arises from

an intrinsic property of the Fresnel zone plate and its cousins: The focal length is so long at short wavelengths that the telescope must be distributed between two spacecraft separated by 0.1–1 km. With the optic on one and the detector on the other, the relative positions of the two spacecraft must be controlled to millimeter accuracy. The technology for such precision formation flying is developing rapidly. Thus, a coronal microscale observatory that implements a classical optical concept using state-of-the-art technology may solve a solar mystery that has endured for the better part of a century.

REFERENCES

- 1. E. W. Cliver, W. F. Dietrich, J. Space Weather Space Clim. 3, A31 (2013).
- R. Ramaty, R. G. Stone, eds., High Energy Phenomena on the Sun, NASA (1973).
- 3. J. A. Klimchuk, Sol. Phys. 234, 41 (2006).
- 4. T. R. Rimmele et al., Sol. Phys. 295, 172 (2020).
- 5. S. Tomczyk et al., J. Geophys. Res. Space Phys. 121, 7470 (2016).
- S. Gosain et al., in *Ground-Based and Airborne Instrumentation for Astronomy III*, C. J. Evans, L. Simard, H. Takami, eds., SPIE (2018), p. 107024H.
- T. S. Bastian, in Solar and Space Weather Radiophysics: Current Status and Future Developments, D. E. Gary, C. U. Keller, eds., Springer (2004), p. 47.
- 8. D. E. Gary et al., https://arxiv.org/abs/2210.10827.
- 9. G. E. Brückner et al., Sol. Phys. 162, 357 (1995).
- D. Rabin, in The WSPC Handbook of Astronomical Instrumentation, vol. 3, A. M. Moore, ed., World Scientific (2021), p. 339.

GET MORE AT PHYSICSTODAY. ORG Our digital-only offerings include webinars, whitepapers, exclusive news stories, and interesting commentaries—all with the same broad coverage as the magazine. Sign up to be alerted when new, exclusive content appears online! PHYSICSTODAY.ORG/SIGNUP