Chapter 9
The Extended Solar Atmosphere

This chapter discusses the Sun's outer atmosphere, the million-degree corona, which expands away from the Sun in fast, uniform winds and slow, gusty ones. Solar flares and coronal mass ejections are also reviewed together with the space-weather effects of these solar explosions on the Earth and nearby space. A complete, in-depth treatment of all of these topics, with numerous references, can be found in Lang, Kenneth R., *The Sun From Space*, *Second Edition*, Heidelberg: Springer-Verlag 2009.

9.1 Hot, Volatile, Magnetized Gas

9.1.1 The Million-Degree Solar Corona

The apparent edge of the visible solar disk, the photosphere, is illusory, for a hot, transparent atmosphere envelops it, extending all the way to the Earth and beyond. This unseen atmosphere is more rarefied than the best vacuum on Earth, and so tenuous that we see right through it.

The diaphanous outer atmosphere of the Sun includes – from its deepest part outward – the underlying photosphere, from the Greek word *photos* for "light"; the thin chromosphere, from the Greek word *chromos* for "color"; and the extended *corona* from the Latin word for "crown." We can observe the chromosphere and corona during a total solar eclipse, when the Moon blocks out the intense light of the underlying photosphere (Fig. 9.1).

Because of their very low densities and high temperatures, the chromosphere and corona produce bright spectral features called emission lines. Atoms and ions in a hot tenuous gas produce such emission features, heated to incandescence and shining at precisely the same wavelengths as the dark absorption lines produced by the same substance in the cooler photosphere. The corona's emission lines provided the initial evidence that it is hundreds of times hotter than the underlying photosphere.

K. R. Lang, *Essential Astrophysics*, Undergraduate Lecture Notes in Physics,
DOI: 10.1007/978-3-642-35963-7_9, © Springer-Verlag Berlin Heidelberg 2013

Fig. 9.1 Eclipse corona The million-degree solar atmosphere, known as the corona, was seen around the shadowed disk of the Moon during the solar eclipse on 11 July 1991. The electrically charged gas was concentrated by magnetic fields into numerous fine rays as well as larger helmet streamers. The expanding corona envelops the Earth and all the other planets. (Courtesy of HAO/NCAR.)

The corona's emission lines were first observed during total eclipses of the Sun (Table 9.1). The intense green line, at a wavelength of 530.3 nm, for example, was first observed during the solar eclipse of August 7, 1869, and for decades attributed to a previously unknown substance dubbed coronium (Young 1869). About 70 years later, it was attributed to emission of iron ions, denoted by Fe XIV, by Walter Grotrian (1890–1954) of Potsdam and Bengt Edlén (1906–1933), a Swedish astronomer who specialized in spectroscopy (Grotrian 1934, 1939; Edlén 1941, 1945). These ions are iron atoms missing 13 of their 26 electrons.

The reason it took so long to identify the coronal emission lines is that no one realized the corona was so hot and also because such spectral features can arise only in the very tenuous corona. They are "forbidden transitions" that do not occur in terrestrial circumstances where collisions between atoms keep them from happening even in the best vacuum.

Iron must be at a temperature of a few million K for atomic collisions to remove so many electrons from the atoms. Edlén provided additional evidence for this hot temperature from the observed widths of the emission lines (Edlén 1941). Elements move at a faster speed in a hotter gas, broadening the observed spectral features as well as producing them. The million-degree temperature of the corona was subsequently confirmed by observations of the Sun's radio radiation (Pawsey 1946) and intense x-ray radiation.

Table 9.1 Strong forbidden emission lines in the visible light of the Sun's low corona

Wavelength (nm)	Ion	Name	Wavelength (nm)	Ion
338.8	Fe XIII		670.2	Ni XV
423.2	Ni XII		789.2	Fe XI
530.3	Fe XIV	Green line	802.4	Ni XV
569.4	Ca XV	Yellow line	1074.7	Fe XIII
637.4	Fe X	Red line	1079.8	Fe XIII

[a] Adapted from Edlén (1941) and Swings (1943). The symbols Ca, Fe and Ni denote, respectively, Calcium, Iron and Nickel. Subtract one from the Roman numeral to obtain the number of missing electrons. Thus, the ion Fe XIII is an iron atom missing 12 electrons. The wavelength is in units of nanometers, or 1 nm = 10^{-9} m. Astronomers have often used the Ångström unit of wavelength, where 1 Ångström = 1 Å = 0.1 nm × 10^{-10} m

Example: **The million-degree corona**

The identification of an emission line of the solar corona with Fe XIV indicated that the coronal gas would have to be very hot. We can estimate how hot that would be by equating the thermal energy of the gas, $3kT/2$, at temperature T, to the ionization potential of Fe XIV, which is 235.04 eV; this is the amount of energy needed to remove so many electrons from the iron atom. Using $k = 1.38065 \times 10^{-23}$ J K^{-1} for the Boltzmann constant and the conversion of 1 eV = 1.602×10^{-19} J, we obtain $T = 2 \times$ ionization potential$/(3k) \approx 1.82 \times 10^{6}$ K, or about 2 million K.

Withbroe and Noyes (1977) provided a review of mass and energy flow in the solar chromosphere and corona. Aschwanden (2006) discussed the physics of the solar corona, and Ashwanden, Poland and Rabin (2001) and Lang (2009) have provided reviews of modern observations of the corona.

Because of its high temperature, the corona emits most of its energy and its most intense radiation as x-rays. The x-rays can be used to image the hot corona all across the Sun's face with high spatial and temporal resolution. This is because the Sun's visible photosphere, being so much cooler, produces negligible x-ray radiation and appears dark under the million-degree corona. Since the Sun's x-ray radiation is absorbed totally in the Earth's atmosphere, it must be observed with telescopes lofted into space by rockets or in satellites.

Modern spacecraft obtain full-disk images of the corona at soft x-ray and extreme ultraviolet wavelengths in lines of ionized iron, Fe XVII, Fe IX, Fe X, Fe XIII, Fe XV, Fe XVI and one line of ionized helium, He II; these are the permitted lines emitted by ionized atoms, sensitive to temperatures from 60,000 to 4.0×10^{6} K (Table 9.2), and not the forbidden lines detected at visible wavelengths.

Close inspection of the Sun's x-ray radiation shows that the star is in constant turmoil, driven by intense, variable magnetic fields. This magnetism is responsible for dark sunspots that temporarily mark the visible face of the Sun.

Table 9.2 Prominent soft x-ray and extreme ultraviolet emission lines from the Sun's low corona and transition region[a]

Wavelength (nanometers)	Emitting ion	Formation temperature (kelvin)
1.70	Iron, Fe XVII	4,000,000
1.90	Oxygen, O VIII	3,100,000
2.16	Oxygen, O VII	2,000,000
3.37	Carbon, C VI	1,300,000
17.11	Iron, Fe IX	630,000
17.45	Iron, Fe X	1,000,000
18.40	Oxygen, O VI	320,000
19.51	Iron, Fe XII	1,400,000
28.42	Iron, Fe XV	2,100,000
30.38	Helium, He II	60,000
33.54	Iron, Fe XVI	2,500,000
33.61	Iron, Fe XVI	2,500,000
46.52	Neon, Ne VII	630,000
60.98	Magnesium, Mg X	1,300,000
155.0	Carbon, C IV	126,000

[a] Subtract one from the Roman numeral to get the number of missing electrons. The wavelengths are in nanometers, abbreviated nm, where 1 nm = 10^{-9} m. Astronomers sometimes use the Ångström unit of wavelength, abbreviated Å, where 1 Å = 10^{-10} meters = 0.1 nm

9.1.2 Varying Sunspots and Ever-Changing Magnetic Fields

The solar corona is permeated by magnetic fields that are generated inside the Sun and rise up through the photosphere into the overlying atmosphere. The strongest magnetism protrudes to blemish the visible Sun with dark, Earth-sized sunspots (Fig. 9.2), which were seen by the unaided human eye up to 3,000 years ago – don't do it, staring at the Sun could burn your eyes.

In the early 20th century, the American astronomer George Ellery Hale (1868–1938) first used the Zeeman effect to show that sunspots are regions of intense magnetism, thousands of times stronger than the Earth's magnetic field (Hale 1908a, b). The intense sunspot magnetism acts as both a valve and a refrigerator, choking off the outward flow of heat and energy from the solar interior and keeping the sunspots cooler and darker than their surroundings.

The strong magnetism exerts a pressure that tends to push apart the magnetic fields; however, by using helioseismology to look under the photosphere astronomers have discovered that flowing material pushes against the magnetic fields of sunspots, holding them in place.

Because the Sun's magnetism is forever changing and is never still, the sunspots are temporary, with lifetimes ranging from hours to months. Moreover, the total number of sunspots varies periodically, from a maximum to a minimum and back to a maximum, in about 11 years (Fig. 9.3). Samuel Heinrich Schwabe (1789–1875), an amateur astronomer in Dessau, Germany, discovered this periodic variation in the mid-nineteenth century (Schwabe 1844). At the maximum in the

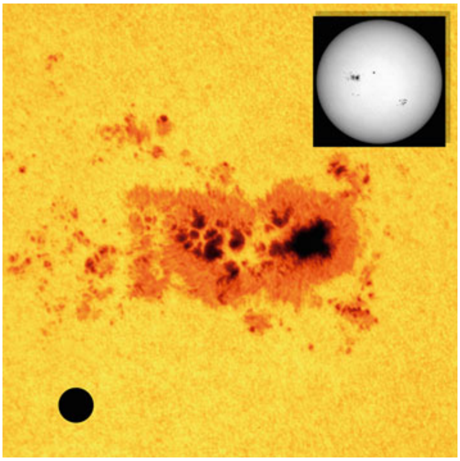

Fig. 9.2 Sunspot group Intense magnetic fields emerge from the interior of the Sun through the Sun's visible disk, the photosphere, producing groups of sunspots. The sunspots appear dark because they are slightly cooler than the surrounding photosphere gas. This composite image was taken in white light; that is, in all of the colors combined. The enlarged image shows the biggest sunspot group, which is about 12 times larger than the Earth, the size of which is denoted by the black spot (*lower left*). (Courtesy of *SOHO*/ESA/NASA.)

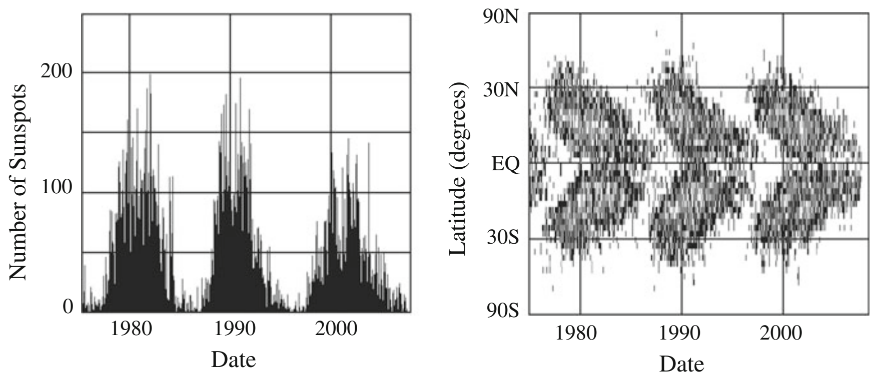

Fig. 9.3 Solar magnetic activity cycle The 11 year solar cycle of magnetic activity is plotted from 1975 to 2007. Both the numbers of sunspots (*left*) and the positions of sunspots (*right*) wax and wane in cycles that peak every 11 years. Similar 11 year cycles have been observed for more than a century. At the beginning of each cycle, the first sunspots appear at about 30° solar latitude and then migrate to 0° solar latitude, at the solar equator (EQ), when the cycle ends. This plot of the changing positions of sunspots resembles the wings of a butterfly, and therefore has been called the *butterfly diagram*. The cycles overlap with spots from a new cycle appearing at high latitudes while the spots from the old cycle persist in the equatorial regions. The solar latitude is the angular distance from the plane of the Sun's equator, which is very close to the plane of the Earth's orbit about the Sun, called the *ecliptic*. (Courtesy of David Hathaway/NASA/MSFC.)

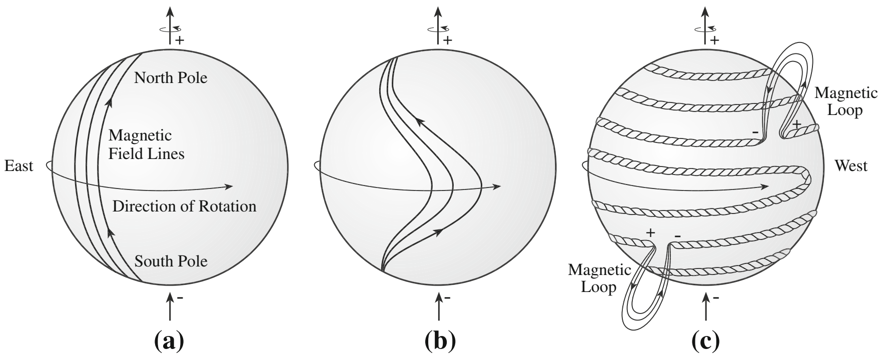

Fig. 9.4 Winding up the field A model for generating the changing location, orientation, and polarity of the sunspot magnetic fields. Initially, the magnetic field is supposed to be the dipolar field seen at the poles of the Sun (*left*). The internal magnetic fields then run just below the photosphere from the Sun's South to North Pole. As time proceeds, the highly conductive, rotating material inside the Sun carries the magnetic field along and winds it up. Because the equatorial regions rotate at a faster rate than the polar regions, the internal magnetic fields are stretched out and wrapped around the Sun's center, becoming concentrated and twisted together like ropes (*middle* and *right*). With increasing strength, the submerged magnetism becomes buoyant, rises and penetrates the visible solar disk – the photosphere – creating magnetic loops and bipolar sunspots that are formed in two belts, one each in the northern and southern hemisphere (*right*). The simplified model shows only two magnetic loops, but many of them are created at about the same time. [Adapted from Horace W. Babcock, "The topology of the Sun's magnetic field, and the 22 year cycle," *Astrophysical Journal* 133, 572–587 (1961).]

sunspot cycle, there may be 100 or more spots on the visible hemisphere of the Sun at one time; at sunspot minimum, very few are seen, and, for periods as long as a month or more, none can be found. The locations where sunspots emerge and disappear also vary over the 11 year sunspot cycle, from mid-latitudes on the Sun to the solar equator (see Fig. 9.3).

The American astronomer Horace W. Babcock (1912–2003) devised a conceptually simple model for the varying sunspots (Babcock 1961). Working with his son, Harold, he had shown that the Sun has a general dipolar magnetic field of about 10^{-4} tesla, usually limited to high solar latitudes near the solar poles (Babcock and Babcock 1955). His dynamo theory begins at sunspot minimum with a global, dipolar magnetic field that runs inside the Sun from south to north, or from pole to pole. Uneven, or differential, rotation – in which the equatorial regions rotate faster than the polar ones – shears the electrically conducting gases of the interior. As a result, the entrained magnetic fields are stretched out and squeezed together. The magnetism is coiled, bunched, and amplified as it is wrapped around the inside of the Sun. The surrounding gas buoys up the concentrated magnetism, and eventually the magnetic fields become strong enough to rise up to the photosphere and break through it in belts of bipolar sunspot pairs (Fig. 9.4).

The initial dipolar magnetic field is twisted into a submerged, ring-shaped field running parallel to the solar equator, or east to west. There are two buried magnetic fields, one in the northern hemisphere and one in the southern hemisphere, but oppositely directed, which bubble up at mid-latitudes to spawn two belts of sunspots, symmetrically placed on each side of the equator.

As the 11 year cycle progresses toward maximum activity, the internal magnetic field is wound increasingly tighter by the shearing action of differential rotation. The two sunspot belts slowly migrate toward the solar equator, where the sunspots in the two hemispheres tend to merge.

Diffusion and poleward flows sweep the remnant magnetism into streams, each dominated by a single magnetic polarity, that slowly wind their way from the low- and mid-latitude belts to the Sun's poles. By sunspot minimum, the continued poleward transport of their debris may form a global dipole with reversed polarity. The north and south poles switch magnetic direction or polarity at the next sunspot minimum. When the Sun's magnetic flip is considered, we see that it takes two activity cycles, or about 22 years, for the overall magnetic polarity to return to where it began. The internal magnetism then has readjusted to its submerged dipolar form, and the magnetic cycle begins again.

9.1.3 Coronal Loops

Magnetic fields are described by lines of force, like those joining the opposite poles of a bar magnet. The direction of the lines of force and the orientation of the magnetic fields can be inferred from the polarization of the spectral lines that have been split by the Zeeman effect. Magnetic-field lines pointing out of the Sun have positive magnetic polarity, whereas inward-directed fields have negative polarity.

Sunspots usually appear in adjacent pairs or other close groupings of opposite magnetic polarity (Hale 1919). Invisible magnetic arches loop between these oppositely directed magnetic regions, often emerging from a sunspot with one polarity and reentering a neighboring sunspot of opposite polarity. Although they remain unseen in optically visible sunlight, these coronal loops shine brightly in x-ray and extreme ultraviolet images of the Sun taken with telescopes in space (Fig. 9.5). Because this radiation is absorbed in our atmosphere, such images cannot be obtained from the ground. Material is concentrated to higher densities and temperatures within these loops, so they emit this invisible radiation more intensely than their surroundings. This intense x-ray and extreme ultraviolet emission thus outlines the magnetic shape and structure of the Sun's outer atmosphere, indicating that the corona is stitched together by bright, thin magnetized loops.

The magnetized atmosphere in, around, and above bipolar sunspot groups is called a solar active region. Active regions are places of concentrated, enhanced magnetic fields, sufficiently large and strong to stand out from the magnetically weaker areas. These disturbed regions are prone to awesome explosions, marking a location of extreme unrest on the Sun.

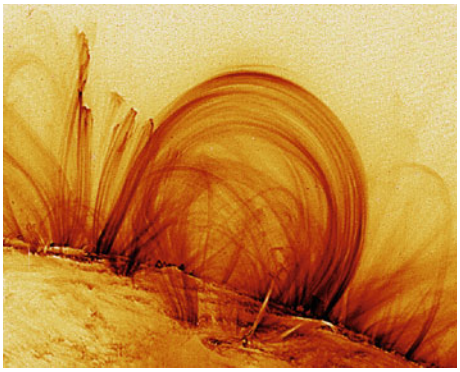

Fig. 9.5 Magnetic loops made visible An electrified, million-degree gas, known as *plasma*, is channeled by magnetic fields into bright thin loops. The magnetized loops stretch up to 500,000 km from the visible solar disk, spanning up to 40 times the diameter of planet Earth. The magnetic loops are seen in the extreme ultraviolet radiation of eight and nine times ionized iron, denoted Fe IX and Fe X, formed at a temperature of about 1.0 million K. The hot plasma is heated at the bases of loops near the place where their legs emerge from and return to the photosphere. Bright loops with a broad range of lengths all have a fine thread-like substructure with widths as small as the telescope resolution of 1 s of arc, or 725 km at the Sun. This image was taken with the *Transition Region And Coronal Explorer* (*TRACE*) spacecraft. [Courtesy of the *TRACE* consortium, LMSAL and NASA; *TRACE* is a mission of the Stanford-Lockheed Institute for Space Research, a joint program of the Lockheed-Martin Solar and Astrophysics Laboratory (LMSAL), and Stanford's Solar Observatories Group.]

Example: **Coronal loops in solar active regions**

Intense magnetic fields of strength $B \approx 0.03$ tesla confine a hot ionized gas in solar active regions within coronal loops. The gas pressure $P_g = N_e kT$ of the hot electrons, of number density N_e, is just equal to the magnetic pressure $P_B = B^2/(2\mu_0)$ required to confine the hot plasma when $N_e \approx 2.6 \times 10^{19}$ m^{-3}, assuming a temperature of $T = 1$ million or 10^6 K and using a Boltzmann constant $k = 1.38065 \times 10^{-23}$ J K^{-1} and the permeability of free space $\mu_o = 4\pi \times 10^{-7}$ N A^{-2} $= 1.2566 \times 10^{-6}$ N A^{-2}. In other words, a million-degree gas with an electron density less than 10^{19} m^{-3} will be constrained by these coronal loops.

If these loops expel their electrons during a solar flare, the electrons will be accelerated to higher energies by the magnetic interaction that triggers and powers the flare (see Sect. 9.3). The energy of the electrons might be comparable to that of x-rays, with $E = 30$ keV $= 4.80 \times 10^{-15}$ J. For a coronal loop of radius R the volume can be approximated as $4\pi R^3/3$, and for a loop radius comparable to the size of the Earth, or $R = 6.378 \times 10^6$ m, the total flare energy released will be $E_f = 4\pi R^3 N_e E/3 \approx 1.25 \times 10^{26}$ J.

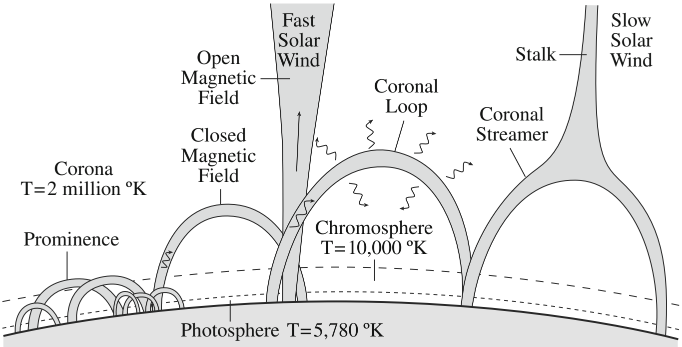

Fig. 9.6 Coronal loops The corona is stitched together with ubiquitous coronal loops that are created when upwelling magnetic fields generated inside the Sun push through the visible solar disk – the photosphere – into the overlying, invisible chromosphere and corona. These closed magnetic structures are anchored in the photosphere at foot points of opposite magnetic polarity. Coronal loops can be filled with hot gas that shines brightly at extreme ultraviolet and x-ray wavelengths. Driven by motions in the underlying photosphere and below, the coronal loops twist, rise, shear, and interact, releasing magnetic energy that can heat the solar corona and power intense solar flares or coronal mass ejections. Large coronal loops are found in the bulb-like base of coronal streamers, whose long, thin stalks extend out into space. Magnetic fields anchored in the photosphere at one end can also be carried by the solar wind into interplanetary space, resulting in open magnetic fields and a channel for the fast solar wind

The number of active regions, with their bipolar sunspots and coronal loops, varies in step with the sunspot cycle, peaking at sunspot maximum when they dominate the structure of the inner corona. At sunspot minimum, the active regions are largely absent and the strength of the extreme-ultraviolet and x-ray emission of the corona is greatly reduced. Because most forms of solar activity are magnetic in origin, the sunspot cycle also is called the solar cycle of magnetic activity.

Unlike the Earth, magnetism on the Sun does not consist of only one simple dipole; it contains numerous interlooped pairs of opposite magnetic polarity. Powerful magnetism, spawned deep inside the Sun, threads its way through the solar atmosphere, creating a dramatic, ubiquitous, and ever-changing panorama of coronal loops (Fig. 9.6).

Throughout the solar atmosphere, a dynamic tension is set up between the gas pressure of the charged particles and the pressure of the magnetic field (Focus 9.1). In the photosphere and convective zone, the gas pressure dominates the magnetic pressure, allowing the magnetic field to be carried around by the moving gas. Because the churning gases are ionized and hence electrically conductive, they sweep the magnetic field along (Fig. 9.7). The situation is reversed in the low corona within active regions, where hot ionized particles are confined within coronal loops (Fig. 9.8).

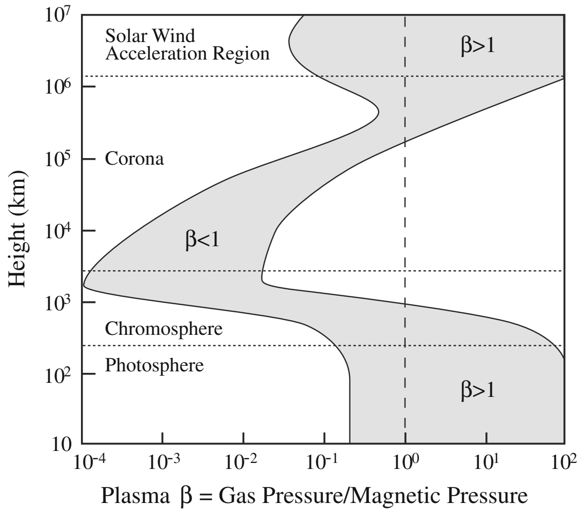

Fig. 9.7 Gas and magnetic pressure The ratio of gas to magnetic pressure, denoted by the symbol β, is plotted as a function of height above the photosphere. The magnetic pressure is greater than the gas pressure in the low corona, where β is less than 1, and magnetic fields determine the structure of the corona. Farther out, the gas pressure can exceed the magnetic pressure, which permits the solar wind to carry the Sun's magnetic field into interplanetary space. In the photosphere, below the corona and chromosphere, the gas pressure also exceeds the magnetic pressure, and the moving gas carries around magnetic fields. [Adapted from Allen (2001).]

Focus 9.1 Magnetic pressure and gas pressure

A magnetic field tends to restrain a collection of electrons and protons, called plasma, while the plasma exerts a pressure that opposes this field. The magnetic pressure is an energy density associated with the magnetic field. The magnetic pressure, P_B, produced by a magnetic field transverse to its direction is given by:

$$P_B = \frac{B^2}{2\mu_0}, \tag{9.1}$$

for a magnetic field of strength B in tesla, where the permeability of free space $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2} = 1.2566×10^{-6} N A^{-2}. As expected, a stronger magnetic field applies a greater restraining pressure.

Hot plasma generates a gas pressure, P_G, owing to the motions of its particles. The ideal gas law describes it:

$$P_G = N\,k\,T, \tag{9.2}$$

where N is the particle number density, $k = 1.38065 \times 10^{-23}$ J K^{-1} is the Boltzmann constant, and T is the temperature. Hotter particles move faster and create greater pressure to oppose the magnetic field, and denser plasma also results in greater gas pressure.

The two kinds of pressure compete for control of the solar atmosphere. In the low solar corona, strong magnetic fields in active regions hold the hot, dense electrified gas within coronal loops. The magnetic and gas pressures become equal for a magnetic field, B, given by:

$$B = [(2\mu_0 k)NT]^{1/2} = [3.48 \times 10^{-29} NT]^{1/2} \text{ tesla.} \tag{9.3}$$

If a coronal loop contains a hot, dense plasma with $N = 10^{17}$ electrons per cubic meter and $T = 10^6$ K, the magnetic field must be stronger than $B \approx 0.002$ tesla to restrain the plasma. By way of comparison, the magnetic field strength at the Earth's equator is 0.00003 tesla, or at least 100 times weaker than the magnetic field in some coronal loops.

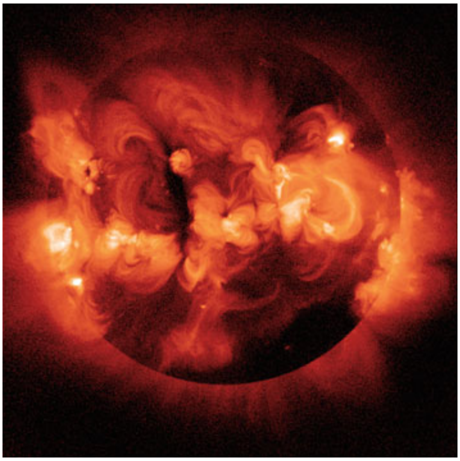

Fig. 9.8 The Sun in x-rays Ionized gases at a temperature of a few million K produce the bright glow seen in this x-ray image of the Sun. It shows magnetic coronal loops that thread the corona and hold the hot gases in place. The brightest features are called active regions and correspond to the sites of the most intense magnetic field strength. The Soft X-ray Telescope (SXT) aboard the Japanese *Yohkoh* satellite recorded this image of the Sun's corona on 1 Feb 1992, near a maximum of the 11 year cycle of solar magnetic activity. Subsequent SXT images, taken about five years later near activity minimum, show a remarkable dimming of the corona when the active regions associated with sunspots have almost disappeared, and the Sun's magnetic field has changed from a complex structure to a simpler configuration. (Courtesy of NASA /ISAS/LMSAL/NAO Japan, University of Tokyo.)

The coronal magnetic fields emerge from underneath the photosphere where they are rooted, and they are continually displaced and replaced by convective motions just below the photosphere. As a result, the corona has no permanent features and it is never still, quiet, or inactive. It is always in a continued state of metamorphosis.

9.1.4 What Heats the Corona?

The visible solar disk, the photosphere, is closer to the Sun's center than the million-degree corona, but the photosphere is several hundred times cooler, with a temperature of 5,780 K. This temperature difference is unexpected because energy should not flow from the cooler photosphere to the hotter corona any more than water should flow uphill. It violates the second law of thermodynamics, which states that heat cannot be continuously transferred from a cooler body to a warmer one without doing work.

We know that visible sunlight cannot resolve the heating paradox. Radiation from the photosphere does not go into the corona; it goes through the corona. There is so little material in the corona that it is transparent to almost all of the photosphere's radiation. Therefore sunlight passes right through the corona without depositing substantial quantities of energy into it, traveling out to warm the Earth and to also keep the photosphere cool.

So, radiation cannot resolve the heating paradox. We must look for alternate sources of energy, and they are related to either moving gases or the magnetic fields in the photosphere and below. Unlike radiation, either the kinetic energy of moving material or the magnetic energy released by magnetic fields can flow from cold to hot regions, keeping the corona hot.

In 1948–1949, astronomers in Germany, the United States, and France independently proposed that sound waves generated in the turbulent convective zone might heat the overlying atmosphere (Biermann 1948; Schwarzschild 1948; Schatzman 1949). The sound waves would accelerate and strengthen as they travel outward through the increasingly rarefied, overlying solar atmosphere, until supersonic shocks are created, dissipating energy and heating the gas.

Although observations from the eighth *Orbiting Solar Observatory*, abbreviated *OSO* 8, showed that sound waves do not transport significant amounts of energy into the corona, these measurements indicated that the sounds might warm the chromosphere to 10,000 K, or roughly twice the temperature of the underlying photosphere (Athay and White 1978, 1979; Bruner 1981). Modern observations indicate that even though the majority of sound waves generated in the convective zone are reflected back into the solar interior at the photosphere, a small percentage of them do manage to slip through the photosphere along inclined magnetic fields, forming shocks that heat the low chromosphere and create numerous short-lived spicules there. This method of chromosphere heating is generally consistent with the fact that other stars with outer convective zones have chromospheres, while stars that have no convective zones do not exhibit a detectable chromosphere.

For coronal heating, other kinds of waves must be considered, and a likely candidate is magnetic waves that can propagate into the corona and carry energy into it. The Sun's ever-changing coronal magnetic fields are always being jostled, twisted, and stirred around by motions deep within the Sun where the magnetism originates. A tension acts to resist the motions and pull the disturbed magnetism back, generating waves that propagate along magnetic fields, somewhat like a vibrating string. These waves do not form shocks, and once generated they can propagate for large distances, directing their energy along open magnetic fields into the overlying corona.

Such waves are now called Alfvén waves after Hannes Alfvén (1908–1995) who first described them mathematically (Alfvén 1942a, b, Sect. 5.6), and argued that they might heat the corona (Alfvén 1947). Ronald G. Giovanelli (1915–1984), Jack H. Piddington (1910–1997), and Donald E. Osterbrock (1924–2007) subsequently discussed the heating of the chromosphere and corona by Alfvén waves (Giovanelli 1949; Piddington 1956; Osterbrock 1961). An important ongoing controversy is whether or not Alfvén waves propagating through the corona dissipate sufficient energy to heat it. The waves have been detected far from the Sun, suggesting that they might heat the distant corona by traveling along open magnetic fields (Belcher 1969; Cranmer and van Ballegooijen 2005).

Strong, interacting magnetic fields also play a role in heating closed magnetic regions closer to the Sun, within the coronal loops. After all, the hottest and densest material in the low corona is located where the magnetic field is strongest, usually within active regions above sunspots. Moreover, observations from solar spacecraft indicate that the entire magnetic flux in the so-called quiet solar atmosphere outside solar active regions is replenished every 15–40 h (Schrijver and Title 2003).

Motions down inside the convective zone twist and stretch the overlying magnetic fields, slowly building up their energy, and magnetic loops of all sizes always are being pushed up into the solar corona from below. When oppositely directed coronal loops are pressed together, they can merge and join at the place where they touch, releasing their energy to heat the corona. The magnetic fields then reform or reconnect in new magnetic orientations, so this method of coronal heating is termed magnetic reconnection.

Thus, magnetic fields seem to play a fundamental role in channeling, storing, and transforming energy into heat, supplying it on different timescales and sending it to various structures. When the magnetic geometry does not change, the magnetism plays a passive role, guiding the flow of charged particles, heat, and waves along the field lines. And when the magnetic configuration changes, the magnetism can play an active role by triggering instabilities and releasing stored magnetic energy through merging and reconnection of closed magnetic field lines.

9.1.5 Coronal Holes

In contrast to the dense, bright areas, the corona also contains less dense regions called coronal holes. These so-called holes have so little material in them that they appear as large dark areas on x-ray or extreme-ultraviolet images, seemingly devoid of radiation (also see Fig. 9.8).

The coronal holes are neither constant nor permanent; they appear, evolve, and die away in periods ranging from a few weeks to several months, continuously changing in content, shape, and form.

At times of low solar activity, near the minimum of the Sun's 11 year magnetic activity cycle, coronal holes cover the north and south polar caps of the Sun. During more active periods, closer to the cycle maximum, the large coronal holes at the poles shrink and even disappear, and smaller coronal holes appear at all solar latitudes – even at or near the solar equator.

The rarefied coronal holes are not completely empty. The normally constraining magnetic forces relax and open up in the coronal holes to allow an unencumbered outward flow of electrically charged particles and magnetic fields into interplanetary space, keeping the coronal hole's density low and expelling a relentless high-speed wind.

9.2 The Sun's Varying Winds

9.2.1 The Expanding Sun Envelops the Earth

The Sun's radiation is not all that passes through the space between the planets. It is filled with electrons, protons, and magnetic fields emanating from the Sun in a ceaseless flow. These unseen particles and fields form a perpetual solar wind that extends all the way to the Earth and far beyond. It was inferred from comet tails, suggested by theoretical considerations, and fully confirmed by direct measurements from spacecraft in the early 1960s (Focus 9.2). So the space between the planets is not completely empty; it contains the Sun's winds that stream out radially in all directions from the Sun.

Focus 9.2 Discovery of the solar wind

The notion that something is always being expelled from the Sun first arose from observations of comet tails. Comets can appear unexpectedly almost anywhere in the sky, moving in every possible direction, but with tails that always point away from the Sun. A comet therefore travels headfirst when approaching the Sun and tail first when departing from it. Ancient Chinese astronomers concluded that the Sun must have a *chi*, or "life force", that blows away the comet tails. In the early 1600s, the German astronomer

Johannes Kepler (1571–1630) proposed that solar radiation pushes the comet tails away from the Sun.

When a comet is tossed into the inner solar system, the dirty ice on its surface is vaporized, sometimes forming two kinds of tails that always point generally away from the Sun rather than toward it. One is a yellow tail of dust, which can litter the comet's curved path. The dust is pushed away from the Sun by the pressure of sunlight. The other tail is colored electric blue, shining in the light of ionized particles. The ions in comet tails always stream along straight paths away from the Sun.

The existence of the solar wind of charged particles was suggested from observations of comet ion tails in the mid 20th century. The German astronomer, Ludwig Biermann (1907–1986) noticed, in the 1950s, that the ions in a comet's tail move with velocities many times higher than could be caused by the weak pressure of sunlight, and proposed that a continued flow of electrically-charged particles pours out of the Sun at all times and in all directions, accelerating the ions to high speeds and pushing them away from the Sun in straight ion tails (Biermann 1951, 1957).

Eugene N. Parker (1927–) of the University of Chicago showed how such a relentless flow might work, dubbing it the *solar wind* (Parker 1958). It would naturally result from the expansion of the Sun's million-degree atmosphere, the corona. He also demonstrated how a magnetic field would be pulled into interplanetary space from the rotating Sun, attaining a spiral shape (Parker 1958, 1960, 1963). The expansion begins slowly near the Sun, where the solar gravity is the strongest. Then the expanding corona accelerates outward into space until the winds break away from the Sun, and eventually they cruise along at the roughly constant and supersonic velocities needed to account for the acceleration of comet tails. This creates a strong, persistent solar wind, forever blowing throughout the solar system.

The first direct measurements of the solar wind's corpuscular, or particle, content were made by a group of Soviet scientists led by Konstantin I. Gringauz (1918–1993), using four ion traps aboard the *Lunik 2* spacecraft launched to the Moon on September 12, 1959. In the following year, Gringauz reported that the maximum current in all four ion traps corresponded to a solar wind flux of 2 million million (2×10^{12}) ions (presumably protons) per square meter per second (Gringauz 1961). This is in rough accord with all subsequent measurements.

All reasonable doubt concerning the existence of the solar wind was removed by measurements made on board NASA's *Mariner 2*, launched on August 27, 1962. Marcia Neugebauer (1932–) and Conway W. Snyder of the Jet Propulsion Laboratory used more than one hundred days of *Mariner 2* data, obtained as the spacecraft traveled to Venus, to show that charged particles are continuously emanating from the Sun, for at least as long as instruments on *Mariner 2* observed them (Neugebauer and Snyder 1966; Neugebauer 1997). It also unexpectedly indicated that the solar wind has a slow and a fast component. The slow one moves at a speed of 300–400 km s^{-1}; the fast one travels at twice that speed.

The solar wind flux determined by Neugebauer and Snyder was in good agreement with the values measured with the ion traps on *Lunik 2*. The average wind ion number density was shown to be 5 million (5×10^6) protons per cubic meter near the distance of the Earth from the Sun. We now know that such a low density close to the Earth's orbit is a natural consequence of the wind's expansion into an ever-greater volume, but that variable wind components can gust with higher densities.

In the low corona, strong magnetic fields constrain the hot ionized gas within coronal loops. But further out in the corona, the magnetic fields decrease in strength and cannot restrain the outward flow of the million-degree gas; it also flows out unencumbered from the open magnetic fields in coronal holes.

Example: **Hot enough to break away from the Sun's gravity**

The corona is fully ionized, with a temperature of several million K and consisting of electrons and protons. We can determine if these particles are hot enough to escape from the Sun's gravity by equating their thermal velocity $(3kT/m)^{1/2}$ to the escape velocity of the Sun $(2GM_\odot/R_\odot)^{1/2}$, to obtain (Sects. 3.2, 5.2, 5.3):

$$T = \frac{2GmM_\odot}{3kR_\odot} \approx 9.2 \times 10^{33}\, m \text{ K}, \tag{9.4}$$

where the Sun's mass $M_\odot = 1.989 \times 10^{30}$ kg and radius $R_\odot = 6.955 \times 10^8$ m, the Boltzmann constant $k = 1.38065 \times 10^{-23}$ J K^{-1} and the constant of gravitation $G = 6.674 \times 10^{-11}$ N m^2 kg^{-2}. The temperature is denoted by T and the particle mass is designated by m.

For a proton of mass $m_P = 1.6726 \times 10^{-27}$ kg, the temperature required for escape is $T \approx 1.5 \times 10^7$ K, while for the electron of mass $m_e = 9.1094 \times 10^{-31}$ kg it is $T \approx 8.4 \times 10^3$ K. So the electrons in the million-degree solar corona would have no problem escaping from the Sun, and the protons would require just a little extra push to put them out where the solar gravity is somewhat diminished.

The solar gale brushes past the planets, wraps itself around the Earth, and carries the Sun's corona out to interstellar space. This radial, supersonic outflow creates a huge bubble of plasma, with the Sun at its center and the planets inside; this is called the *heliosphere*, from *Helios* the "God of the Sun" in Greek mythology.

9.2.2 Properties of the Solar Wind

The million-degree corona is so hot that it cannot stand still. Indeed, the solar wind consists of an overflow corona, which is too hot to be entirely constrained by the Sun's inward gravitational pull and perpetually moves out into surrounding space. The hot gas creates an outward pressure that tends to oppose the inward pull of the Sun's gravity. At great distances, where the solar gravity weakens, the hot protons and electrons overcome the Sun's gravity and continue to accelerate, like water overflowing a dam. So, the solar corona is really the visible, inner base of the solar wind, and the solar wind is simply the hot corona expanding into interplanetary space.

The Sun's continuous wind travels with two main velocities. There is a fast, uniform wind that blows at a speed of about 750 km s^{-1}, and a variable, gusty slow wind that moves about half as fast. Both winds are supersonic, moving at least 10 times faster than the speed of sound in the solar wind. The Sun's wind also rushes on with little reduction in speed because there is almost nothing out there to slow it down. Both the fast and slow winds from the Sun are much more tenuous, hotter, and faster than any wind on the Earth.

Example: **The supersonic solar wind**

The speed of sound, c_s, in the solar wind can be determined from the equation (Laplace 1816, Sect. 5.8):

$$c_S = \left(\frac{\gamma kT}{\bar{m}}\right)^{\frac{1}{2}}, \tag{9.5}$$

where the adiabatic index $\gamma = 5/3$ for a fully ionized gas like the solar wind, the Boltzmann constant $k = 1.38065 \times 10^{-23}$ J K^{-1}, and the mean molecular mass $\bar{m} = \mu \times$ u where the mean molecular weight $\mu = 0.5$ for the solar wind (fully ionized almost entirely composed of protons), and the atomic mass unit u $= 1.66054 \times 10^{-27}$ kg. For a temperature of $T = 1.2 \times 10^5$ K for solar-wind protons near the Earth's orbit, this equation gives $c_s \approx 5.8 \times 10^4$ m $s^{-1} \approx 58$ km s^{-1}. By way of comparison, the slow and fast solar winds have respective velocities of about 375 and 750 km s^{-1}, which indicates that the solar wind is everywhere supersonic, or moving at speeds faster than sound.

Table 9.3 Mean values of solar-wind parameters at the Earth's orbit[a]

Parameter	Mean value
Particle density, N	$N \approx 10^7$ m^{-3} (5 electrons and 5 protons per cubic centimeter)
Velocity, V	Fast wind $V \approx 750$ km s^{-1}
	Slow wind $V \approx 375$ km s^{-1}
	Average $V \approx 600$ km s^{-1}
Mass density, ρ	$\rho = 10^{-20}$ kg m^{-3} (protons)
Flux, F	$F \approx 6 \times 10^{12}$ particles m^{-2} s^{-1}
Temperature, T	$T \approx$ 120,000 K (protons) to 140,000 K (electrons)
Particle thermal energy, kT	$kT \approx 2 \times 10^{-18}$ J $\approx$ 12 eV
Proton kinetic energy, 0.5 m_pV^2	0.5 $m_pV^2 \approx 10^{-16}$ J $\approx$ 1,000 eV = 1 keV
Particle thermal energy density	$NkT \approx 10^{-11}$J m^{-3}
Proton kinetic energy density	0.25 $N\, m_p\, V^2 \approx 10^{-9}$ J m^{-3}
Radial magnetic field, B_r	$B_r = 2.5 \times 10^{-9}$ T = 2.5 nT = 2.5 $\times 10^{-5}$ G
Alfvén velocity, V_A	$V_A = 32$ km s^{-1}
Sound speed, c_s	$c_s \approx 50$ km s^{-1}

[a] These solar-wind parameters are at the mean distance of the Earth from the Sun, or at one astronomical unit, 1 AU, where 1 AU = 1.496 $\times 10^{11}$ m. The Boltzmann constant k =1.38 $\times 10^{-23}$ J K^{-1} relates temperature and thermal energy. The proton mass $m_p = 1.67 \times 10^{-27}$ kg

Because the electrified wind material is an excellent conductor of heat, the temperature falls off only gradually with distance from the Sun, reaching between 120,000 K and 140,000 K at the Earth's distance. The tenuous wind has been diluted to a rarefied plasma by the time it reaches the Earth's distance from the Sun, where there are approximately 5 million electrons and 5 million protons per cubic meter of solar wind. The density of the solar wind is so low that there are not enough particles to heat astronauts that might venture outside their spacecraft.

Because it is wrapped into a spiral, the interplanetary magnetic field strength falls off linearly with distance from the Sun, in contrast to the solar wind number density that decreases more rapidly, as the inverse cube of the distance, as it fills a larger volume.

Physical properties of the solar wind at the Earth's distance from the Sun are listed in Table 9.3.

At a mean speed of about 600 km s^{-1}, the flux of solar wind particles is far greater than anything else in nearby space. Between one and ten million million (10^{12}–10^{13}) particles in the solar wind cross every square meter of space each second. That flux far surpasses the flux of more energetic cosmic rays that enter our atmosphere, with the abundant protons arriving with a flux of only 640 protons per square meter per second. The peak local energy density of cosmic rays is about one million electron volts per cubic meter, or about one ten thousandth (10^{-4}) the kinetic energy density of solar wind protons.

Moreover, interplanetary magnetic fields act as a barrier to electrically charged cosmic rays that are coming from the depths of space, preventing them from reaching the Earth. During the maximum in the solar cycle, stronger solar magnetic fields are carried out into interplanetary space by the Sun's wind, deflecting more cosmic rays. Less extensive interplanetary magnetism, during a minimum in the 11 year cycle of magnetic activity, lowers the barrier to the cosmic particles and allows more of them to arrive at Earth. This unexpected anti-correlation between solar activity and cosmic rays arriving at Earth is often called the Forbush effect after its discovery by Scott Forbush (1904–1984) in the early 1950s (Forbush 1950). It was explained by Peter Meyer (1920–2002), Eugene Parker and John Simpson (1916–2000) who proposed that enhanced interplanetary magnetism near the maximum in the 11 year solar activity cycle deflects cosmic rays from their Earth-bound paths (Meyer et al. 1956).

Because the Sun is blowing itself away continuously, we might imagine that it would eventually vanish from view after expelling all of its substance into space. Every second, the solar wind carries about a billion kilograms, or 1 million tons, of the Sun into surrounding space. That seems significant, but in 4.6 billion years the solar wind has only carried away about 0.0001, or one ten thousandth, of the Sun's mass $M_{\odot} = 1.989 \times 10^{30}$ kg. Moreover, that is about three times less than the amount of mass turned into energy during this time by nuclear reactions near the center of the Sun (Focus 9.3).

Focus 9.3 Mass loss from the Sun

The Sun is relentlessly expelling mass into the solar wind, most of it in protons with a mass $m_P = 1.6726 \times 10^{-27}$ kg. Using measurements at the Earth, located at a mean distance from the Sun of 1 AU $= 1.496 \times 10^{11}$ m, we can estimate the total mass, ΔM_{SW} lost from the Sun in the solar wind every second:

$$\Delta M_{SW} = 4\pi m_P N_P V(\mathrm{AU})^2 = 1.41 \times 10^9 \text{ kg}, \tag{9.6}$$

where the number density of protons in the solar wind at the Earth's distance is $N_P = 5 \times 10^6$ m^{-3} and we have assumed an average wind velocity of $V = 600$ km s^{-1} $= 6 \times 10^5$ m s^{-1}. The total mass M_{SW} lost at this rate over the past 4.6 billion years, using 1 year $= 3.156 \times 10^7$ s, is:

$$M_{SW} = \Delta M_{SW} \times 4.6 \times 10^9 \times 3.156 \times 10^7 = 2.05 \times 10^{26} \text{ kg} \approx 0.0001 M_{\odot}, \tag{9.7}$$

where $M_{\odot} = 1.989 \times 10^{30}$ kg denotes the Sun's current mass.

About three times more mass is consumed every second by nuclear reactions that make the Sun shine. For just one fusion of four protons into one helium nucleus the mass lost is $\Delta m = 0.007\ (4m_P)$ with an energy release $\Delta E = \Delta m c^2$, for a velocity of light $c = 2.9989 \times 10^8$ m s^{-1}. Since these nuclear reactions are supplying the Sun's absolute luminosity $L_{\odot} = 3.828 \times 10^{26}$ J s^{-1}, the mass loss ΔM_{NR} in one second is:

$$\Delta M_{NR} = \Delta m \left(\frac{L_{\odot}}{\Delta E} \right) = \frac{L_{\odot}}{c^2} \approx 4.26 \times 10^9 \text{ kg}, \tag{9.8}$$

and over the past 4.6 billion years, assuming the Sun has always been shining at the same rate, the total mass M_{NR} lost by nuclear reactions is:

$$M_{NR} = 4.6 \times 10^9 \times 3.156 \times 10^7 \times \Delta M_{NR} \approx 6.2 \times 10^{26} \text{ kg} \approx 0.0003\, M_\odot. \tag{9.9}$$

This mass is carried off by the radiation that makes the Sun shine.

9.2.3 Where Do the Two Solar Winds Come From?

Instruments aboard spacecraft have detected two solar winds with different physical properties. There is a fast wind that moves at a speed of about 750 km s^{-1} and a slow wind that blows at about half that speed. The high-speed wind is steady and uniform, whereas the slow-speed wind is variable and gusty.

The two solar winds do not blow uniformly from all places on the Sun; instead they depend on solar latitude. The spatial distribution of the two types of winds also depends on the Sun's magnetic field configuration, which varies dramatically with the 11 year solar-activity cycle.

As suggested by Sir William I. Axford (1933–2010), the steady, uniform, high-speed wind emanates from magnetically open configurations in the corona (Axford 1985). The open magnetic fields in coronal holes provide a conduit for the fast wind. In contrast the slow wind – which is filamentary and transient – involves the intermittent release of material from previously closed magnetic regions, so the slow wind may not be treated as an equilibrium flow in a steady state.

The distribution of the open and closed magnetic regions on the Sun, and therefore the places of origin for the two solar winds, depends on the 11 year cycle of solar magnetic activity.

Near activity minimum, the high-speed wind moves out of the open magnetic fields in large coronal holes located at the Sun's polar regions. A slow, gusty, and variable wind then moves away from closed magnetic regions near the Sun's equator (McComas et al. 2000).

The simple, bimodal distribution of fast and slow wind structures disappears near the maximum in the 11 year solar activity cycle. The large polar coronal holes then shrink and even disappear and smaller coronal holes appear at all solar latitudes. A chaotic and complex mixture of varying solar wind flows therefore is found at all solar latitudes near activity maximum (McComas et al. 2003). The slow winds still seem to be associated with closed magnetic structures, such as active regions, whereas the fast winds rush out of the interior of coronal holes all over the Sun. Solar active regions, with their explosive behavior, provide a noticeable third source for the solar winds near the activity-cycle maximum.

9.2.4 Where Does the Solar Wind End?

How far does the solar wind extend, and where does its influence end? The solar wind carves out a cavity in the interstellar medium known as the *heliosphere*. Zurbuchen (2007) has reviewed the coupling of the Sun and the heliosphere; also see Lang (2009).

Since the solar wind is weakened by expansion, thinning out as it moves into a greater volume, it eventually becomes too dispersed to repel interstellar forces. The winds are no longer dense or powerful enough to withstand the pressure of gas and magnetic fields coursing between the stars. The radius of this celestial standoff distance, in which the pressure of the solar wind falls to a value comparable to the interstellar pressure, has been estimated at about 100 AU, or one hundred times the mean distance between the Earth and the Sun.

The radius of the heliosphere can be estimated by determining the standoff distance, or stagnation point, in which the ram pressure, P_W, of the solar wind falls to a value comparable to the interstellar pressure, P_I. As the wind flows outward, its velocity remains nearly constant, while its density decreases as the inverse square of the distance. The dynamic pressure of the solar wind therefore also falls off as the square of the distance, and we can use the solar-wind properties at the Earth's distance of 1 AU to infer the pressure, P_{WS}, at the stagnation-point distance, R_S. Equating this to the interstellar pressure we have:

$$P_{WS} = P_{1AU} \times \left(\frac{1AU}{R_S}\right)^2 = \left(N_{1AU}\, V_{1AU}^2\right) \times \left(\frac{1AU}{R_S}\right)^2 = P_I, \tag{9.10}$$

where the number density of the solar wind near the Earth is about $N_{1AU} = 5$ million particles per cubic meter and the velocity there is about $V_{1AU} = 500$ km per second.

To determine the distance to the edge of the solar system, R_S, we also need to know the interstellar pressure, which is the sum of the thermal pressure, the dynamic pressure, and the magnetic pressure in the local interstellar medium. Its estimated value results in $R_S = 100$ AU or more, well beyond the orbits of the major planets.

The termination shock at the edge of the solar wind has now been measured from the *Voyager 1* and *2* spacecraft (Decker et al. 2005; Stone et al. 2005, 2008).

Instruments aboard the twin *Voyager 1* and *2* spacecraft, launched in 1977 and now cruising far beyond the outermost planets, have approached this edge of the solar system from different directions. *Voyager 1* is moving in the northern hemisphere of the heliosphere and *Voyager 2* in the southern hemisphere. *Voyager 1* crossed the termination shock of the supersonic flow of the solar wind on December 16, 2004 at a distance of 94 times the mean distance between the Earth and the Sun, or at 94 AU from the Sun. At this distance, the spacecraft's instruments recorded a sudden increase in the strength of the magnetic field carried by the solar wind, as expected when the solar wind slows down and its particles pile up at the termination shock.

Voyager 2 crossed the termination shock on August 30, 2007 at a distance of 84 AU from the Sun. It appears that there is a significant north/south asymmetry in the heliosphere, likely due to the direction of the local interstellar magnetic field.

Both *Voyager 1* and *2* have therefore now crossed into the vast, turbulent heliosheath, the region where the interstellar gas and solar wind interact, due to the reflection and deflection of the solar-wind ions by the magnetized wind beyond the heliosheath. In technical terms, the solar-wind ions in the heliosheath are deflected by magnetosonic waves reflecting off of the heliopause, causing the ions to flow parallel to the termination shock toward the heliotail.

Both *Voyager* spacecraft are equipped with plutonium power sources expected to last until at least 2020 and perhaps 2025. So they ought to eventually measure the heliopause at the outer edge of the heliosheath. It is the place where interstellar space begins.

In the meantime, the *Interstellar Boundary EXplorer*, abbreviated *IBEX*, was launched on 19 Oct 2009. Instruments on this spacecraft, which operates in Earth orbit, detect neutral, or unionized, atoms coming from the termination shock and the boundary between the solar wind and interstellar space.

9.3 Explosions on the Sun

9.3.1 Solar Flares

Suddenly, without warning, a part of the Sun explodes, creating a solar flare. Some of them are the biggest explosions in the solar system, releasing energy of ten million, billion, billion joule, or 10^{25} J, in just 100 s. This is comparable in strength to 20 million nuclear bombs exploding simultaneously, and each with energy of 100 Megatons of TNT.

A substantial fraction of the flare energy goes into accelerating electrons and protons to nearly the speed of light. Some of these high-energy particles are hurled into the Sun, briefly raising the temperature of Earth-sized regions of the Sun to more than 10 million K. Other accelerated particles are tossed out into interplanetary space and emit intense radio and x-ray radiation.

The short-lived solar flares unleash their energy in the vicinity of sunspots, covering just a few tenths of a percent of the solar disk. These incredible explosive outbursts become more frequent and violent when the number of sunspots is greatest; several solar flares can be observed on a busy day near the maximum of the sunspot cycle. However, they are not caused directly by sunspots; solar flares instead are powered by magnetic changes in the corona above sunspots.

Although it emits very large amounts of energy, a solar flare usually releases less than 1/1,000th of the total energy radiated by the Sun every second, so they are only minor perturbations in the combined colors, or white light, of the Sun. The first record of a solar flare detected on the visible solar disk therefore did not occur until the mid-nineteenth century – on September 1, 1859, when the English astronomers Richard Christopher Carrington (1826–1875) and Richard Hodgson (1804–1872) independently noticed one (Carrington 1860; Hodgson 1860).

A new perspective, which demonstrated the frequent occurrence of solar explosions, was made possible when flares were observed in the red Balmer alpha emission line of hydrogen at 656.3 nm, originating in the chromosphere. This wasn't possible until the early 20th century, after the invention of instruments that isolated the red hydrogen-alpha emission from the Sun's intense visible light at adjacent wavelengths in the solar spectrum. Such observations showed that at the chromospheric level in the solar atmosphere a solar flare consists in simplest form as two extended, parallel flare ribbons. But a fundamental understanding of the physical processes responsible for solar flares had to wait until they were detected at invisible radio wavelengths from the ground and in x-rays from space.

During World War II (1939–1945) it was discovered that sudden, intense radio outbursts from the Sun, associated with solar flares, could interfere with radio communications and radar systems. Soon after the war ended, J. Paul Wild's (1923–2008) group of Australian radio astronomers used swept frequency radio receivers to show that some flares eject particles at about half the speed of light, or about 1.5×10^8 m s^{-1}, while others moved at about 1/100th this speed or at about 10^6 m s^{-1}, and were attributed to shock waves. Wild and Smerd (1972) and Wild et al. (1963) provided early summaries of observations of solar radio bursts. Bastian et al. (1998) provided a more recent review of the radio emission from solar flares.

Example: **Watching flare-accelerated electrons and shock waves**

When a flare-associated disturbance, such as an electron beam or a shock wave, moves through the coronal plasma, the local electrons are displaced with respect to the protons, which are more massive than the electrons. The electrical attraction between the electrons and protons pulls the electrons back in the opposite direction, and an oscillation is set up at the plasma frequency, ν_p, given by (Sect. 5.5):

$$\nu_P = \left[\frac{e^2 N_e}{4\pi^2 \varepsilon_0 m_e}\right]^{1/2} = 8.98 N_e^{1/2} \text{ Hz}, \tag{9.11}$$

for an electron density N_e, where the electron charge $e = 1.6022 \times 10^{-19}$ coulomb, the electron mass $m_{\mathrm{e}} = 9.1094 \times 10^{-31}$ kg, the permittivity of free space is $\varepsilon_0 = 8.8542 \times 10^{-12}$ F m^{-1}, and $\pi = 3.14159$. Note that the wavelength, λ_p, of radiation at the plasma frequency is given by $\lambda_p \nu_p = c = 2.9979 \times 10^8$ m s^{-1}, the speed of light.

Low in the solar corona, where $N_e \approx 10^{14}$ m^{-3}, the plasma frequency is about 90 MHz, at radio frequencies, or radio wavelengths of about 3.3 m, which can be observed from the ground. Since the plasma frequency decreases with the diminishing coronal electron density at greater distances from the Sun, where the corona occupies an increasing volume of space, beams or shocks sent out from a solar flare excite progressively lower plasma frequencies. Near the Earth, for example, $N_e \approx 5 \times 10^6 \text{ m}^{-3}$ and the plasma frequency is about 20 kHz corresponding to a wavelength of 15 km.

When an electron density model of the solar atmosphere is used, the emission frequency can be related to height, and combined with the time delays observed at successively lower frequencies, to obtain the outward velocity of the moving disturbance. For a type III radio burst, an average of about 0.4 c is determined. The slower drift associated with type II radio bursts suggests an outward motion at about 1,000 km s^{-1} and has been attributed to shock waves.

Radiation from low-frequency plasma oscillations, between 100 kHz and 10 MHz, is absorbed in the Earth's atmosphere, but they have been observed for decades using instruments aboard spacecraft. These observations can be used to track the spiral magnetic field that guides the flare electrons as they move out through the increasingly rarefied coronal plasma.

The bulk of radiation from high-temperature solar flares is not emitted as radio waves but instead at extreme ultraviolet and x-ray wavelengths, where they can briefly outshine the entire Sun. This radiation is absorbed in the Earth's atmosphere; therefore, astronomers have observed it from outer space, beginning with primitive instruments aboard balloons or sounding rockets and continuing with increasingly sophisticated telescopes in many satellites, including the *Yohkoh*, *Ulysses*, *Wind*, *SOHO*, *ACE*, *TRACE*, *Hinode*, *RHESSI*, *STEREO*, and *SDO* spacecraft (Lang 2009).

Why does a solar flare occur? What triggers the instability and suddenly ignites an explosion from magnetic fields that remain unperturbed for long intervals of time? Answers to these questions were proposed in 1960 by the Austrian-born American Thomas Gold (1920–2004), then at Cornell University, and by the Englishman Fred Hoyle (1915–2001), at Cambridge University, when they showed how solar flares could be powered by stressed magnetic loops that interact, dissipating their energy in the corona (Gold and Hoyle 1960). The flares are triggered when magnetized coronal loops are pressed together, driven by motions beneath them, meeting to touch one another and merge.

Magnetic fields have a direction associated with them and if oppositely directed magnetic fields are pushed together, they can interact. When these merging magnetic fields are closed coronal loops, they will break open to release magnetic energy in the form of flare heating and particle acceleration. The magnetic fields are not broken permanently; they simply reconnect to their closed state. For this reason, this merging and coupling is known as *magnetic reconnection* (Sweet 1969; Hirajama 1974).

Benz and Güdel (2010) provided a review of magnetically driven flares on the Sun and other stars, whereas Zweibel and Yamada (2009) reviewed magnetic

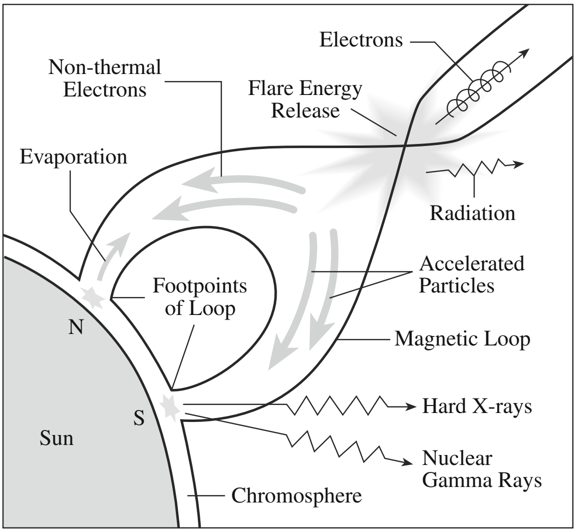

Fig. 9.9 Solar flare model A solar flare is powered by magnetic energy released from a magnetic interaction site above the top of a coronal loop. Electrons are accelerated to high speed during a solar flare, generating a burst of radio energy as well as impulsive loop-top hard X-ray emission. Some of these nonthermal electrons are channeled down the loop and strike the chromosphere at nearly the speed of light, emitting hard X-rays by electron−ion *bremsstrahlung* at the loop footpoints. When beams of accelerated protons enter the dense, lower atmosphere, they cause nuclear reactions that result in gamma-ray spectral lines and energetic neutrons. Material in the chromosphere is heated very quickly and rises into the coronal loop, accompanied by a slow, gradual increase in soft X-ray radiation. This upwelling of heated material is called *chromospheric evaporation* and it occurs in the decay phase of the flare

reconnection in astrophysical plasmas. Bhattacharjee (2004) discussed magnetic reconnection in the Earth's magnetotail and the solar corona.

So, powerful solar flares stem from the interaction of coronal loops. These loops are always moving about, and often are brought into contact by these movements. Magnetic fields coiled up in the solar interior, where the Sun's magnetism is produced, also can bob into the corona to interact with preexisting coronal loops. In either case, the coalescence leads to the rapid release of magnetic energy through magnetic reconnection.

Because flares apparently originate in the low corona and the ubiquitous coronal loops dominate their structure, it's not surprising that solar-flare models involve a coronal loop (Fig. 9.9). Magnetic reconnection triggers the release of magnetic energy just above the loop top, where electrons and protons are accelerated. In less than 1 s, electrons are accelerated to nearly the speed of light, producing intense radio signals. Protons likewise are accelerated to high speeds, and both the electrons and protons are hurled down into the Sun and out into space.

Instruments aboard spacecraft have observed nuclear reactions and the creation of anti-matter during solar flares (Chupp 1984; Share et al. 2004; Hurford et al.

2006). When protons and heavier ions are accelerated to high speed during solar flares, and beamed downward into the Sun, they slam into the dense, lower atmosphere, shattering the nuclei of atoms and sometimes tearing energetic neutrons out of them. Many of these neutrons are eventually captured by ambient, or non-flaring, hydrogen nuclei, the protons, in the photosphere, emitting one of the Sun's strongest gamma-ray lines at 2.223 MeV.

Another strong gamma-ray line emitted during solar flares is the 0.511 MeV pair-annihilation line. Positrons, the anti-matter counterpart of electrons, are released during the decay of radioactive nuclei produced when flare-accelerated protons and heavier nuclei are hurled down into the lower solar atmosphere during a solar flare. The positrons and the electrons annihilate, producing radiation at 0.511 MeV, which is the energy contained in the entire mass of a nonmoving electron.

Example: **Anti-matter created during solar flares**

Positrons, or positive electrons, are the anti-matter particles of electrons, and they are produced during solar flares. The positrons, denoted e^+, are inferred from observations of the pair-annihilation reaction during solar flares. This reaction is written as:

$$e^+ + e^- = \gamma + \gamma, \tag{9.12}$$

where e^- denotes an electron and γ is a gamma-ray photon. The energy, E, of this photon will be equal to the energy taken to completely destroy an electron of rest mass $m_e = 9.1094 \times 10^{-31}$ kg, or $E = m_e\, c^2 = 8.187 \times 10^{-14}$ J = 0.511 MeV, where the speed of light $c = 2.9979 \times 10^8$ m s^{-1}, and 1 MeV = 1.60217×10^{-13} J. Since the positron has the same mass as an electron, both gamma-ray protons have this energy. The 0.511 MeV line has been observed during solar flares using instruments aboard spacecraft.

Because the chromosphere has been heated very rapidly by the accelerated particles that were hurled down into it, that part of the chromosphere explodes, or evaporates, up into the corona to release the excess energy. This process may include the gradual release of energy when the coronal loop relaxes into a more stable configuration during the decay phase of a solar flare.

Example: **Chromospheric evaporation**

Energetic particles are hurled down into the chromosphere during a solar flare, heating it to temperatures of $T = 20$ million, or 2×10^7, K, which is slightly hotter than the center of the Sun. The hot gas flows back up into coronal loops, emitting x-ray radiation during the decay phase of the solar flare. The photon energy of the x-rays is roughly equal to the thermal energy of the gas, or $3kT/2 = 4.14 \times 10^{-16}$ J, where the Boltzmann constant $k = 1.38065 \times 10^{-23}$ J K^{-1}. X-ray astronomers like to specify this energy in keV, where 1 keV $= 1.602 \times 10^{-16}$ J, and chromospheric evaporation is detected at x-ray photon energies of about 2.6 keV.

9.3.2 Coronal Mass Ejections

Coronal mass ejections are gigantic magnetic bubbles that can rush away from the Sun at supersonic speeds, expanding to become larger than the Sun in a few hours (Figs. 9.10, 9.11). They carry about 10^{13} kg or, 10 billion tons, of material out into space, produce intense shock waves, and accelerate vast quantities of energetic particles in interplanetary space. A coronal mass ejection moves through space at a speed of about 400 km s^{-1}, carrying a kinetic energy of about 10^{24} J, which is comparable to the explosive energy of a large solar flare. When directed at the Earth, a coronal mass ejection arrives at the planet about four days after being ejected from the Sun (Focus 9.4).

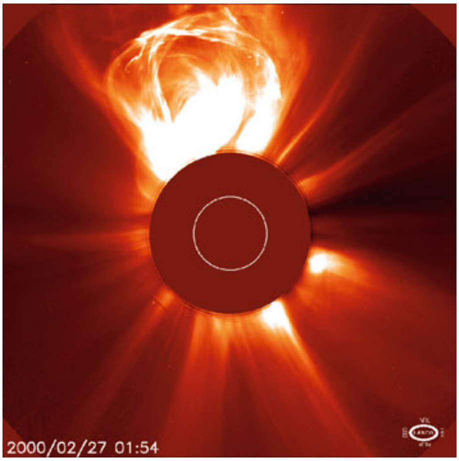

Fig. 9.10 Coronal mass ejection A huge coronal mass ejection is seen in this image, taken on 27 Feb 2000 with a coronagraph on the *SOHO* spacecraft. The white circle denotes the edge of the Sun's visible disk, so this mass ejection is about twice as large as the Sun. The dark area corresponds to the occulting disk of the coronagraph that blocks intense sunlight and permits the overlying solar atmosphere, or corona, to be seen. (Courtesy of the *SOHO* LASCO consortium, *SOHO* is a project of international cooperation between ESA and NASA.)

Focus 9.4 Physical properties of coronal mass ejections

Coronal mass ejections are detected as localized brightness increases in white-light coronagraph images. Integration of the brightness increase, that depends only on the electron density, N_e, permits evaluation of the total mass, M, of the ejection. For a sphere of radius, R, we have:

$$M = 4\pi R^3 N_e m_p / 3, \tag{9.13}$$

where $\pi \approx 3.14159$ and the proton mass $m_p = 1.6726 \times 10^{-27}$ kg. The corona is a fully ionized, predominantly (90 %) hydrogen, plasma, so the number density of protons and electrons are equal, but since the protons are 1,836 times more massive than the electrons, the protons dominate the mass. For a mass ejection with an electron, or proton, density of $N_e = 10^{13}$ m^{-3}, that has grown as large as the Sun, with a radius of $R = 6.955 \times 10^8$ m, this expression gives

$$M \approx 2 \times 10^{13} \text{ kg}, \tag{9.14}$$

or about 20 billion tons.

At the rate of one ejection per day, and 10^{13} kg per ejection, this amounts to a mass flow rate of about 2×10^8 kg s^{-1}, since there are 86,400 s per day.

By way of comparison, the solar wind flux observed in the ecliptic at the orbit of the Earth is about 6×10^{12} protons m^{-2} s^{-1}, or about 10^{-14} kg m^{-2} s^{-1} (Table 9.3). If this flux is typical of that over the entire Sun-centered sphere, with an average Sun-Earth distance of $D = 1$ AU $= 1.496 \times 10^{11}$ m, we can multiply by the sphere's surface area, $4\pi D^2$, to obtain a solar wind mass flow rate of about 2.8×10^9 kg s^{-1}. That's roughly 15 times the mass flow rate from coronal mass ejections, or in other words the coronal mass ejection rate is about 5 % that of the steady, perpetual solar wind, and for just a relatively brief time during the ejection.

The kinetic energy, KE, of a coronal mass ejection with a speed of $V = 400$ km s^{-1} and a mass $M = 10^{13}$ kg is:

$$KE = MV^2/2 \approx 10^{24} \text{ J}. \tag{9.15}$$

This is comparable to the energies of large solar flares that lie between 10^{21} and 10^{25} J.

At a speed of $V = 400$ km s^{-1}, the time, T, to travel from the Sun to the Earth, at an average distance of $D = $ AU $= 1.496 \times 10^{11}$ m is:

$$T = D/V \approx 3.74 \times 10^5 \text{ s} \approx 4.3 \text{ days}, \tag{9.16}$$

where 1 day = 86,400 s. Zhang and Low (2005) have reviewed the hydromagnetic nature of solar coronal mass ejections.

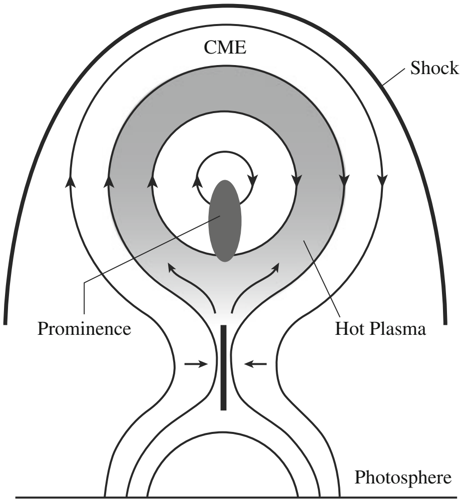

Fig. 9.11 Model of coronal mass ejection A magnetic reconnection takes place at a current sheet (*dark vertical line*) beneath a prominence and above closed magnetic field lines. The coronal mass ejection (CME) traps hot plasma below it (*shaded region*). The solid curve at the *top* is the bow shock driven by the CME. The closed field region above the prominence (*center*) is supposed to become a flux rope in the interplanetary medium. [Adapted from Martens and Kuin (1989).]

The physical size of the mass ejections dwarfs that of solar flares and even the active regions in which flares occur. However, like solar flares, the rate of occurrence of coronal mass ejections varies in step with the 11 year cycle of solar magnetic activity, ballooning out of the corona several times a day during activity maximum. Large coronal mass ejections can occur with or without a solar flare, but they both appear to be powered by the abrupt release of the corona's magnetic energy, with threatening effects for the Earth and nearby space.

9.4 Space Weather

9.4.1 Earth's Protective Magnetosphere

Our planet is immersed within the hot, electrically charged solar wind that blows out from the Sun in all directions and never stops, carrying with it a magnetic field rooted in the Sun. Solar flares and coronal mass ejections create powerful gusts in the Sun's winds, producing space weather – the cosmic equivalent of a terrestrial blizzard or hurricane. Fortunately, we are protected from the full force of this relentless, stormy gale by the Earth's magnetic field.

William Gilbert (1544–1603), physician to Queen Elizabeth I of England, authored a treatise in Latin, with the grand title *De Magnete*, *Magneticisque Corporibus*, *et de Magno Magnete Tellure*, which translated into English is *Concerning Magnetism*, *Magnetic Bodies*, *and the Great Magnet Earth* (Gilbert 1600). In this work, which is still available in an English version, Gilbert showed that the Earth is itself a great magnet, which explains the orientation of compass needles. It is as if there was a colossal bar magnet at the center of the Earth.

At the Equator, the two ends of a compass needle point north or south, toward the Earth's magnetic poles. At each magnetic pole, the needle stands upright, pointing into or out of the ground. In between, at intermediate latitudes, the compass needle points north or south with a downward dip of one end but not vertically as at a pole. Because the geographic poles are located near the magnetic ones, a compass needle is aligned in the north–south direction. We usually put an arrow on the north end of the needle; therefore, an arrowed compass points north.

We can describe the Earth's magnetism by invisible magnetic field lines, which orient compass needles. These lines of magnetic force emerge from the south magnetic pole, loop through nearby space and reenter at the north magnetic pole. The lines are close together near the magnetic poles where the magnetic force is strong, and spread out above the Earth's Equator where the magnetism is weaker than at the poles. We cannot see the invisible magnetic field lines, but compass needles point along them, and other instruments can be used to measure their strength.

The magnetic field strength at the Earth's magnetic equator is 0.0000305 tesla, or 0.305×10^{-4} tesla. Measurements of the surface magnetic fields of the Earth show stronger fields near the poles where the magnetic field lines congregate, at roughly twice the strength of the field at the Equator. Although these fields decrease in strength as the inverse cube of the distance, they remain strong enough to divert most of the solar wind around the Earth at a distance far above the atmosphere, thereby protecting humans on the ground from possibly lethal solar particles.

When any charged particle encounters a magnetic field, it must change direction, moving away from or around the magnetism. When the protons and electrons in the gusts or steady flow of the solar wind encounter the Earth's magnetic fields, they are deflected around it, like a rock in a stream or a windshield deflecting air around an automobile.

The Earth's magnetic fields hollow out a protective cavity in the solar wind, which is called the *magnetosphere*, a term coined in 1959 by Austrian-born American Thomas Gold (1920–2004), then at Cornell University (Gold 1959a, b). It is that region surrounding any planet in which its magnetic field dominates the behavior of electrically charged particles, such as electrons, protons, and other ions.

The dipolar (two poles) magnetic configuration applies near the surface of the Earth, but farther out, the magnetic field is distorted by the Sun's perpetual wind. Although it is exceedingly tenuous, the solar wind is powerful enough to mold the outer edges of the Earth's magnetosphere into a changing asymmetric shape (Fig. 9.12).

The solar wind usually bends around the Earth's magnetic field at a distance from the Earth's center of about 10 times the Earth's radius on the dayside that faces the Sun (Focus 9.5). Here, the solar wind pushes the Earth's magnetism in, compressing its outer magnetic boundary and forming a shock wave, shaped like waves that pile up ahead of the bow of a moving ship and resembling the flow of air around a supersonic aircraft. After forming this bow shock, the solar wind is

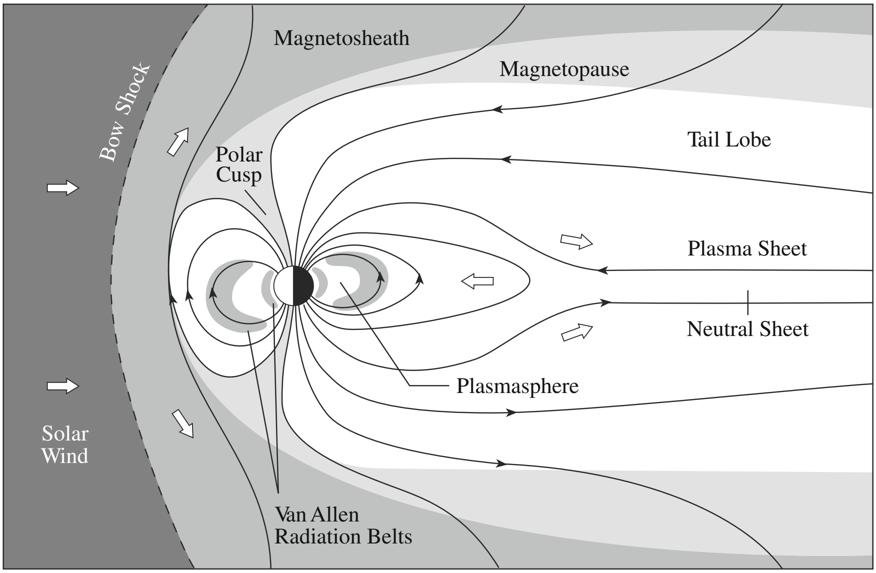

Fig. 9.12 Elements of the magnetosphere The Earth's magnetic field carves out a hollow in the solar wind, creating a protective cavity called the *magnetosphere*. A bow shock forms at about 10 Earth radii on the sunlit side of our planet. The location of the bow shock is highly variable because it is pushed in and out by the gusty solar wind. The magnetopause marks the outer boundary of the magnetosphere, at the place where the solar wind takes control of the motions of charged particles. The solar wind is deflected around the Earth, pulling the terrestrial magnetic field into a long magnetotail on the nightside. Plasma in the solar wind is deflected at the bow shock (*left*), flows along the magnetopause into the magnetic tail (*right*), and then can be injected back toward the Earth within the plasma sheet (*center*). The Earth, its auroras, atmosphere, and ionosphere, and the two Van Allen radiation belts all lie within this magnetic cocoon

deflected around the Earth, pulling the terrestrial magnetic field into a long magnetotail on the night side. Thus, the Earth's magnetosphere is not precisely spherical. It has a bow shock facing the Sun and a long magnetotail in the opposite direction. The term magnetosphere therefore does not refer to form or shape but instead implies a sphere of influence.

Focus 9.5 Planetary magnetospheres

Six planets are known to have magnetospheres. The size of the magnetosphere, on the day side facing the Sun, is determined by the distance, R_{MP}, along the planet-Sun line at which the pressure of the planetary magnetic field balances the dynamic ram pressure of the solar wind.

The ram pressure, P_R, exerted on a stationary body by a gas, or fluid, of mass density, ρ, moving at a velocity, V, is given by:

$$P_R = \rho V^2. \tag{9.17}$$

This expression can also be used for the ram pressure exerted on a body that is moving at this velocity through a gaseous or fluid medium.

For the solar wind, we can use:

$$P_R = m_P N_P V^2, \tag{9.18}$$

where the proton mass $m_P = 1.6726 \times 10^{-27}$ kg, the N_P is the number density of protons at the distance from the Sun that is under consideration, and V is the solar wind velocity at that distance.

The magnetic pressure, P_B, at the surface of the planet is given by:

$$P_B = \frac{B_0^2}{2\mu_0}, \tag{9.19}$$

where B_0 is the equatorial magnetic field strength, $\mu_0 = 4\pi \times 10^{-7}$ N A^{-2} = 1.2566 × 10^{-6} N A^{-2} is the permeability of free space. Since the dipole's magnetic field strength falls off as the cube of the distance from the planet, the magnetic pressure decreases as the sixth power of that distance.

The standoff distance, R_{MP}, from the planet at which the two pressures are equal, or where $P_R = P_B$, therefore occurs when:

$$\frac{R_P^6 B_0^2}{2\mu_0 R_{MP}^6} = m_P N_P V^2, \tag{9.20}$$

where the planet's radius is R_P. Solving for R_{MP} we have:

$$R_{MP} = \left(\frac{B_0^2}{2\mu_0 m_P N_P V^2}\right)^{1/6} R_P. \tag{9.21}$$

Example: Solar-wind bow shock distance for the Earth and Jupiter

When the solar wind encounters the Earth the number density of protons is $N_P \approx 5 \times 10^6$ m^{-3}, and the average solar wind velocity is about $V = 600$ km s^{-1} (see Table 9.3). The equatorial magnetic field strength of the Earth is $B_{0E} = 3 \times 10^{-5}$ tesla. Substituting these numbers into the equation for the standoff point, where the solar wind ram pressure equals the Earth's magnetic pressure, gives $R_{ME} \approx 7R_E$, or seven times the Earth's radius.

The values of R_{MP} for the other planets can be inferred by noting that the solar wind number density, N, falls off with the inverse cube of the distance of the planet from the Sun, while the solar wind velocity remains relatively constant. The giant planets have stronger magnetic fields than the Earth, so the bow shock distances for Jupiter and Saturn, for example, are 42 and 19 times the radius of these planets. The planet Venus has no global dipolar magnetic field, but the solar wind can be diverted around this planet by its thick atmosphere.

The planet Jupiter has an equatorial magnetic field strength of $B_{0J} = 4.28$ tesla, with a bow shock distance of $R_{MJ} = 42\ R_J \approx 3.0 \times 10^9$ m, where the radius of Jupiter is $R_J = 7.15 \times 10^7$ m. So, the magnetosphere of Jupiter is larger than the Sun, whose radius is $R_{\odot} = 6.955 \times 10^8$ m. The solar wind proton density is diminished as it spreads out to a greater volume between the Earth and Jupiter.

The Earth's magnetic shield is so perfect that only 0.1 % of the mass of the solar wind that hits it manages to penetrate inside. Yet, even that small fraction of wind particles has a profound influence on the Earth's nearby environment in space; they create an invisible world of energetic particles and electric currents that flow, swirl, and encircle the Earth.

9.4.2 Trapped Particles

One of the first scientific discoveries of the Space Age was the finding, by James A. Van Allen (1914–2006) and his students, of high-energy electrons and protons that girdle the Earth far above the atmosphere (Van Allen et al. 1959). They move within two belts that encircle the Earth's magnetic equator but do not touch it, like a gigantic, invisible, torus-shaped doughnut.

These regions sometimes are called the inner and outer Van Allen radiation belts. Van Allen used the term "radiation belt" because the charged particles were then known as corpuscular radiation; the nomenclature does not imply either electromagnetic radiation or radioactivity. The radiation belts lie within the inner magnetosphere at distances of 1.5 and 4.5 Earth radii from the center of the Earth, and contain high-speed electrons and protons.

In 1907, about a half-century before the discovery of radiation belts, the Norwegian geophysicist Carl Størmer (1874–1957) showed how electrons and protons could be almost permanently confined and suspended in space by the Earth's dipolar magnetic field (Størmer 1907, 1955). An energetic charged particle moves around the magnetic fields in a spiral path toward one magnetic pole. Its trajectory becomes more tightly coiled in the stronger magnetic fields close to a magnetic pole, where the intense polar fields act like a magnetic mirror, turning the particle around so that it moves back toward the other pole.

Thus, the electrons and protons bounce back and forth between the north and south magnetic poles. It takes about 1 min for an energetic electron to make one trip between the two polar mirror points. The spiraling electrons also drift eastward, completing one trip around the Earth in about a half-hour. There is a similar drift for protons but in the westward direction. The bouncing can continue indefinitely for particles trapped in the Earth's radiation belts, until the particles collide with one another or some external force distorts the magnetic fields.

The problem at the time that Størmer developed his theory was that there was no mechanism known to allow electrically charged particles into the dipolar magnetic field. After all, if electrons and protons cannot leave the magnetic cage, how could they get into it in the first place? The answer is the solar wind. They can arrive via the solar wind and penetrate the Earth's magnetic defense through a temporary opening in it.

The solar wind carries the Sun's magnetic field with it, and the solar magnetism is draped around the magnetosphere when encountering it. As postulated by the English physicist James Dungey (1923–), a solar magnetic field can open up the Earth's magnetic field when the two fields are pointing in opposite directions when they touch (Dungey 1961). With this orientation they can join one another and become linked, similar to how the opposite poles of two toy magnets stick together. The merging process, known as *magnetic reconnection*, can create an opening in the Earth's magnetic field, forming a portal through which the solar particles can flow. Tons of high-energy particles may then flow into the magnetosphere through the opening before it closes again.

9.4.3 Earth's Magnetic Storms

The Earth's magnetic field can be compressed and distorted when a coronal mass ejection arrives at the Earth. The ejections have magnetic field strengths of about 30 nT, or 3×10^{-8} T, so the much stronger terrestrial magnetic field, which is about a thousand times more intense, at 3×10^{-5} T, usually provides good protection from them. However, if the magnetic fields of a coronal mass ejection and of the Earth are pointing in opposite directions when they meet, the two fields become linked, resulting in intense geomagnetic storms that cause compass needles to swing widely.

The flow of currents associated with these great magnetic storms can interfere with electrical power grids here on the Earth, creating voltage surges on long-distance power lines and overheating or melting the windings of transformers. They can send cities into complete darkness, especially in high-latitude regions where the currents are strongest, such as Canada, the northern United States and Scandinavia. This doesn't occur often, perhaps once a year; however, the potential consequences are serious enough to employ early warning systems.

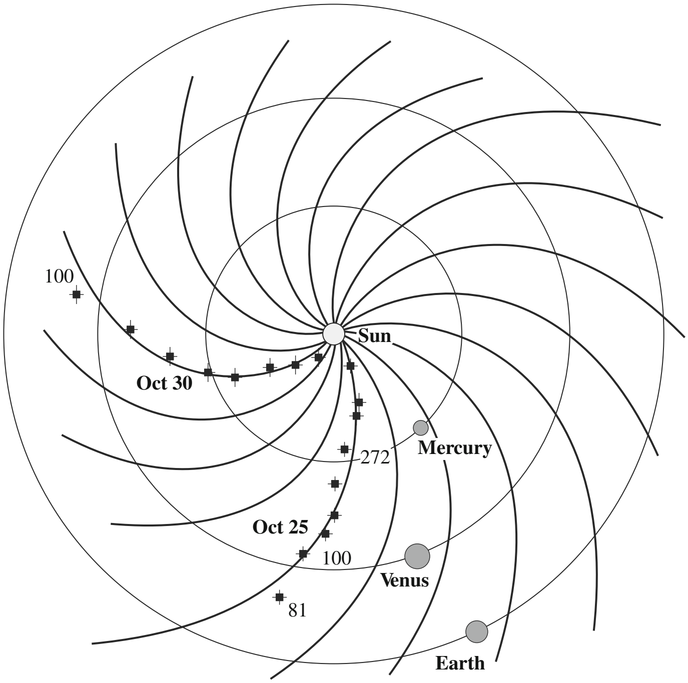

Fig. 9.13 Magnetic spiral The trajectory of flare electrons in interplanetary space as viewed from above the Sun's polar regions using the *Ulysses* spacecraft. The squares and crosses show *Ulysses* radio measurements of Type III radio bursts. As the high-speed electrons move out from the Sun into progressively more tenuous plasma, they excite radiation at successively lower plasma frequencies. The numbers denote the observed frequency in kilohertz (kHz). Because the flaring electrons are forced to follow the interplanetary magnetic field, they do not move in a straight line out from the Sun but instead travel along the spiral pattern of the interplanetary magnetic field, shown by the solid curved lines. The magnetic fields are drawn out into space by the radial solar wind and remain attached at one end to the rotating Sun. The locations of the orbits of Mercury, Venus, and the Earth are shown as circles. (Courtesy of Michael J. Reiner. *Ulysses* is a project of international collaboration between ESA and NASA.)

9.4.4 Solar Explosions Threaten Humans in Outer Space

When directed at our planet or at humans in deep space, both solar flares and coronal mass ejections produce dangerous gusts and squalls in the Sun's winds. Here on the ground, we are shielded from many of the effects by the Earth's atmosphere and magnetic fields, but out in space there can be no protection, and both humans and satellites are vulnerable.

Energetic charged particles generated during a solar flare threaten our planet only if the flare occurs at the right place on the Sun – that is, at one end of the spiral magnetic field that connects the Sun to the Earth. Given this circumstance, when a flare occurs near the west limb and the solar equator, the magnetic spiral guides the high-speed charged particles that can threaten astronauts or satellites. The spiral magnetic pattern, produced when the solar wind carries the rotating Sun's magnetic field into surrounding space, has been confirmed by tracking the radio emission of charged particles thrown out during solar flares, as well as by spacecraft that have sampled the interplanetary magnetism near the Earth (Fig. 9.13).

As suggested by Cornell astrophysicist Thomas Gold (1920–2004), closed magnetic fields can be ejected from the Sun, generating shocks as they move into interplanetary space (Gold 1959a, b). And when a coronal mass ejection travels out into space, it can take the form of a magnetic cloud that moves behind an interplanetary shock (Fig. 9.14). The mass ejection plows into the slower-moving solar wind, driving huge shock waves that cross magnetic field lines and accelerate particles as they go. Following the shocks, the magnetic cloud can remain attached magnetically to the Sun, carrying its looping magnetic fields all the way to the Earth.

Because of their higher mass, it is solar-flare protons rather than electrons that provide the greatest threat to humans in outer space. Such high-speed protons, called solar energetic particles, can endanger the health and even the lives of astronauts when they are in outer space, unprotected by the Earth's magnetic field that deflects the charged particles.

High-energy protons from a solar flare or coronal mass ejection easily can pierce a spacesuit, causing damage to human cells and tissues, and even threaten the life of unprotected astronauts, who venture into space to unload spacecraft cargo, construct a space station or walk on the Moon, which has no global magnetic field or atmosphere, or Mars, which has no magnetosphere and only a thin atmosphere. So, solar astronomers keep careful watch over the Sun during space missions to warn of possible activity occurring in just the wrong place or time.

9.4.5 Disrupting Communication

Eight minutes after an energetic solar flare, a strong blast of x-rays and extreme ultraviolet radiation reaches the Earth and radically alters the structure of the planet's upper atmosphere, known as the ionosphere, altering its ability to reflect radio waves. During even moderately intense flares, long-distance radio communications can be silenced temporarily over the Earth's entire sunlit hemisphere. These radio blackouts are particularly troublesome for the commercial airline industry, which uses radio transmissions for weather, air traffic, and location information. The U.S. Air Force operates a global system of ground-based radio and optical telescopes and taps into the output of national space-borne x-ray telescopes and particle detectors to continuously monitor the Sun for intense flares that might severely disrupt military communications and satellite surveillance.

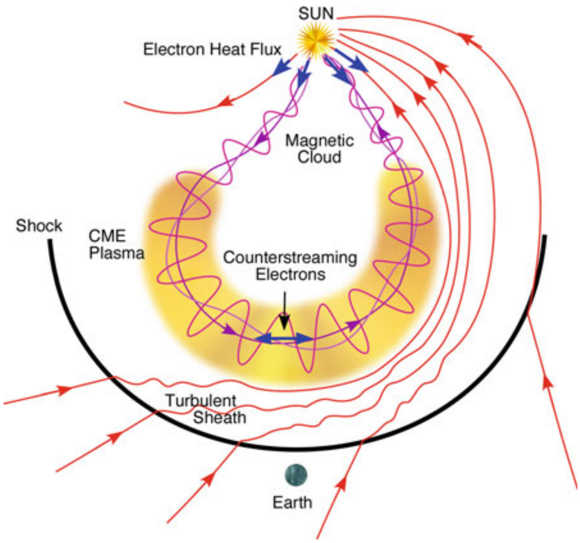

Fig. 9.14 Magnetic cloud When a coronal mass ejection (CME) travels into interplanetary space, it can create a huge magnetic cloud containing beams of electrons that flow in opposite directions within the magnetic loops that are rooted at both ends in the Sun. The magnetic cloud also drives a shock ahead of it. Magnetic clouds are present only in a subset of observed interplanetary coronal mass ejections. (Courtesy of Deborah Eddy and Thomas Zurbuchen.)

Although solar flares do not affect short-wavelength microwave signals that pass right through the ionosphere to communication satellites, solar explosions can destroy the satellites.

9.4.6 Satellites in Danger

More than 1,000 commercial, military, and scientific satellites are now in operation, affecting the lives of millions of people. And the performance and lifetime of all of these satellites are affected by Sun-driven space weather.

Geosynchronous satellites, which orbit the Earth at the same rate that the planet spins, stay above the same place on the Earth to relay and beam down signals used for cellular phones, global positioning systems, and Internet commerce and data transmission. They are endangered by the coronal mass ejections that cause intense geomagnetic storms. These satellites orbit our planet at about 6.6 Earth radii, or about 4,200 km, moving around the Earth once every 24 h. A coronal mass ejection can compress the Earth's protective magnetic fields from their usual location at about 10 Earth radii above the equator to below the satellites' synchronous orbits, exposing them to the full brunt of the gusty solar wind and its charged, energized ingredients.

Other satellites revolve around our planet in closer, low-Earth orbits at altitudes of 300 to 500 km, scanning the air, land, and sea for environmental change, weather forecasting and military reconnaissance. Space weather can increase the atmospheric friction exerted on these satellites, causing their orbits to decay more quickly than expected. The enhanced extreme ultraviolet and x-ray radiation from solar flares heats the atmosphere and causes it to expand; similar or greater effects are caused by coronal mass ejections. The expansion of the terrestrial atmosphere brings higher gas densities to a given altitude, increasing the friction and drag exerted on a satellite, pulling it to a lower altitude, and sometimes causing ground controllers to lose contact with them. Space stations, for example, periodically must be boosted in altitude to a higher orbit to avoid a similar fate.

Infrequent, anomalously large eruptions on the Sun can hurl energetic protons toward the Earth and elsewhere in space. The solar protons can enter a spacecraft, producing erroneous commands and crippling their microelectronics. Such single-event upsets already have destroyed at least one weather satellite and disabled several communications satellites. However, to put the space-weather threat in perspective, only a few commercial satellites have been lost to storms from the Sun out of thousands deployed. The U.S. military builds satellites that can withstand the effects of a nuclear bomb exploded in space.

9.4.7 Forecasting Space Weather

Recognizing our vulnerability, astronomers use telescopes on the ground and in situ particle detectors or remote-sensing telescopes on satellites to carefully monitor the Sun, and government agencies post forecasts that warn of threatening solar activity. This enables evasive action that can reduce disruption or damage to communications, defense and weather satellites, as well as electrical power systems on the ground. Once it is known that a Sun storm is imminent, the launch of manned space flight missions can be postponed, and a walk outside a spacecraft or on the Moon or Mars can be delayed. Airplane pilots can be warned of potential radio communication failures. Operators can power down sensitive electronics on communication and navigation satellites until the danger passes. Utility companies can reduce load in anticipation of trouble on power lines, in that way trading a temporary “brown out” for a potentially disastrous “black out.”

Everyone wants to know how strong a space storm is and when it is going to hit. Like winter storms on the Earth, some of the effects can be predicted days in advance. A coronal mass ejection, for example, arrives at the Earth one to four days after solar astronomers watch it leave the Sun. Solar flares are another matter; as soon as one is observed on the Sun, its radiation and fastest particles have already reached the Earth, taking just 8 min to travel from the Sun. One promising prediction technique is to observe when the magnetism on the Sun has become twisted into a stressed situation, because it then may be about to release a solar flare. Another technique employs helioseismology to look through the Sun and watch active regions develop before they rotate to face the Earth.