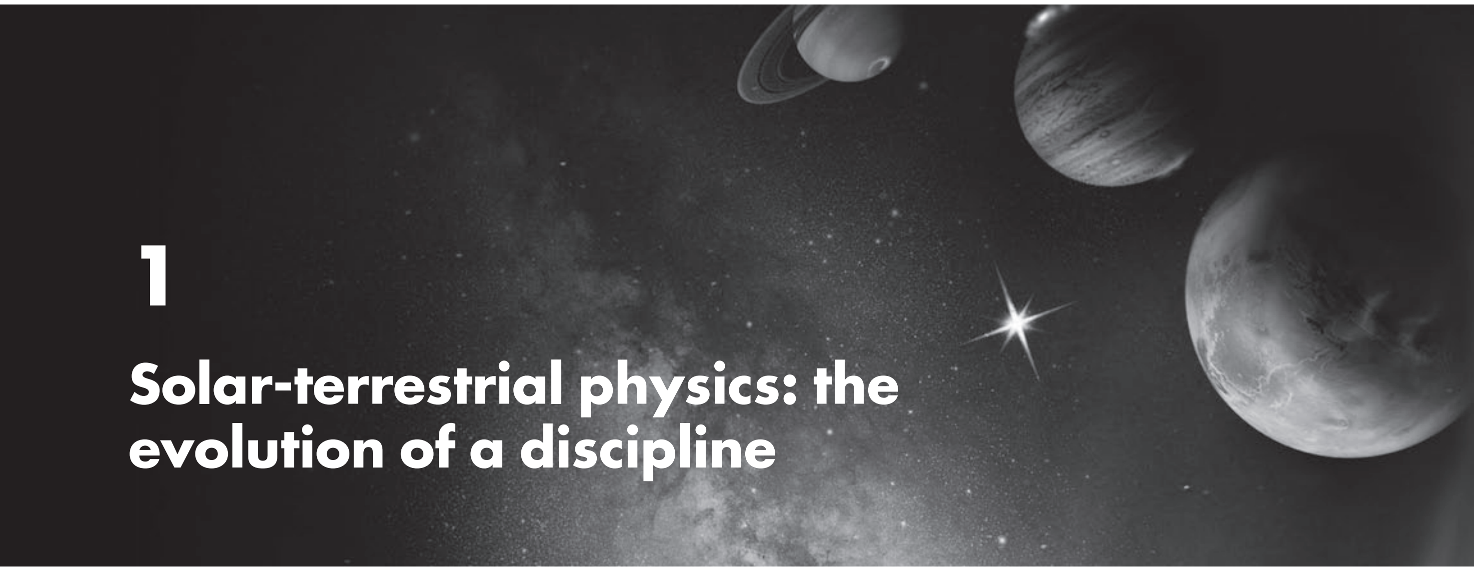

1

Solar-terrestrial physics: the evolution of a discipline

1.1 INTRODUCTION

The physics of the solar-terrestrial environment is governed principally by the interaction of energetic charged particles with electric and magnetic fields in space. Near the Earth, most of these charged particles derive their energy ultimately from the Sun, directly or through the interaction of the solar wind with the Earth's magnetosphere. These interactions are complex and non-linear. The magnetic and electric fields that determine the motion of the charged particles are affected in turn by the motion of these particles. Moreover, the processes that occur on the microscale may affect the behavior of the system on the macroscale. Fortunately, there are approximate methods to treat these systems. We do not have to worry about the motions of the individual particles in most situations.

Some solar-terrestrial research is still carried out on the surface of the Earth with cameras, photometers, spectrometers, magnetometers, and other devices sensitive to the processes occurring high in the upper atmosphere and magnetosphere. However, many processes are invisible to remote sensing and must be studied *in situ*. Thus, the majority of this research is performed using rockets and satellites that enable measurements to be obtained directly in the regions in which the interactions occur. In recent years, these *in situ* data have resulted in explosive growth in our knowledge and understanding of solar-terrestrial processes. Our planetary exploration program has added to this understanding by showing us how such processes operate in quite different settings.

The field of **solar-terrestrial physics** has a long history of investigation, starting well before the advent of satellites and rockets. A convenient way to introduce ourselves to this exotic field, with often non-intuitive, and even counter-intuitive, behavior, is to learn about it in the same sequence as the early pioneers. Thus, we briefly review its historical development to provide context for our later, more physically oriented, presentation of the nature of the processes occurring in the solar-terrestrial environment. We choose to follow a chronological path through solar-terrestrial physics, as that is the way it was revealed to early scientific observers, and this is the way it was originally understood.

1.2 ANCIENT AURORAL SIGHTINGS

The discipline of solar-terrestrial physics began with a growing appreciation of two phenomena: the **auroras** and the variability of the geomagnetic field. Because they can be observed visually, auroras were the first of these phenomena to be recorded. The discovery of the variability of the geomagnetic field had to await the advent of new technology with the invention of the compass.

References to auroras are contained in ancient literature from both the East and West. Chinese literature describes possible auroral sightings, several of which occurred prior to 2000 BC. Several

FIGURE 1.1. Early drawing of an aurora seen on January 12, 1570. (Original print in Crawford Collection, Royal Observatory, Edinburgh.)

passages in the Old Testament appear to have been inspired by auroral sightings. Greek literature includes references to phenomena most likely to have been auroral phenomena. For example, Xenophanes, in the sixth century BC, mentions "moving accumulations of burning clouds." This sequence of discovery may have been influenced by the westward drift of the Earth's tilted **magnetic dipole** that favored auroral displays, first in China, then in Asia Minor, followed by Europe and, today, the eastern United States.

Because the phenomenon was not understood, much fear and superstition surrounded those early sightings of auroras. Figure 1.1, inspired by an auroral display in 1570, illustrates the lack of scientific understanding prevalent at that time. The seventeenth century marked the beginning of scientific theories concerning the origin of the lights in the north. Galileo Galilei, for example, proposed that auroras were caused by air rising out of the Earth's shadow to where it could be illuminated by sunlight. He also appears to have coined the term "aurora borealis," meaning "dawn of the north." At about the same time, Pierre Gassendi, a French mathematician and astronomer, deduced that auroral displays must be occurring at great heights, because they were seen to have the same configuration when observed at places quite remote from one another. His contemporary, René Descartes, is credited with the idea that auroras were caused by reflections from ice crystals in the air at high latitudes. From about 1645 to about 1715, both solar activity and auroral sightings declined, although neither was completely lacking.

Edmund Halley, at the age of 60, after finally observing an auroral display, suggested that auroral phenomena were ordered by the direction of the Earth's magnetic field. In 1731, the French philosopher de Mairan ridiculed the then-popular idea that auroras were a reflection of polar ice and snow, and he also criticized Halley's theory. Instead, he suggested that auroras were connected to the solar atmosphere, and he suspected a connection between the return of sunspots and auroras. Only much later could Halley's ideas and those of de Mairan be reconciled, and studies of geomagnetism and auroras become more firmly linked.

1.3 EARLY MEASUREMENTS OF THE GEOMAGNETIC FIELD

The earliest indication of the existence of the geomagnetic field was the direction-finding capability of the compass. As compasses were improved, more and more was learned about the geomagnetic field. Chinese knowledge that a compass points north or south can be traced to the eleventh century. The encyclopedist Shon-Kau (AD 1030–93) stated that "fortune-tellers rub the point of a needle with the stone of the magnet in order to make it properly indicate the south." In European literature, the earliest mention of the compass and its application to navigation appeared near the end of the twelfth century in *De Untensilibus* and *De Rerum*, two works by Alexander Neekan, a monk of St. Albans, where, coincidentally, much later, one of the authors of this book was born. In the former work, he described the use of the magnetic needle to indicate north and noted that mariners used that means to find their course when the sky was cloudy. In the latter work, he described the needle as being placed on a pivot, a second-generation form of the compass. In neither work did he describe the instrument as a novelty; it was in common use at that time. Official records indicate that, by the fourteenth century, many sailing ships carried compasses. The directions of magnetic north and of geographic or true north differ over most of the globe. The measure of this difference is referred to as the declination. It is not clear when magnetic declination was discovered. However, a letter written in 1544, by Georg Hartmann, vicar of St. Sebald's at Nürnberg, to Duke Albrecht of Prussia showed that he had observed the declination of Rome in 1510 to be 6° east, whereas it was 10° at Nürnberg. Also, it is known that between the years 1538 and 1541, João de Castro made 43 determinations of declination during a voyage along the west coast of India and in the Red Sea.

The geomagnetic field is also inclined to the horizontal. To measure this angle, known as the inclination, one must pivot a needle about a horizontal axis. Georg Hartmann's letter also discussed such an observation, but the angle of inclination was incorrect for his point of observation. William Gilbert ascribed the discovery of the magnetic dip or inclination to an Englishman, Robert Norman, who in 1576 published a work with the title *The newe Attractiue containyng a short discourse of the Magnes or Lodestone, and amongest other his vertues, of a newe discouered secret and subtill propertie, concernyng the Declinyng of the Needle, touched there with onder the plaine of the Horizon. Now first found out by ROBERT NORMAN Hydrographer. Here onto are annexed certaine necessarie rules for the art of Nauigation, by the same R.N. Imprinted at London by John Kyngston, for Richard Ballard, 1581.*

The year 1600 saw the publication of the famous treatise *De Magnete* by William Gilbert, who, in 1601, was appointed chief physician in personal attendance to Queen Elizabeth. This treatise consists of six books containing a total of 115 chapters. The central theme of the book is also the title of Chapter 17, Book 1: "That the globe of the earth is magnetic, a magnet; how in our hands the magnet stone has all the primary forces of the earth, while the earth by the same powers remains constant in a fixed direction in the universe." Figure 1.2 illustrates Gilbert's woodcut, showing the distribution of magnetic inclination or dip over the Earth, and over a small spherical lodestone, which he called a "terrella." Gilbert believed that the terrestrial magnetic field was constant, but it is not. Henry Gellibrand, professor of astronomy at Gresham College, discovered that magnetic declination changed with time, and he published his discovery in a work entitled *A discourse mathematical on the variation of the magneticall needle. Together with its admirable diminuation lately discovered, London, 1635.*

Another early pioneer in the study of geomagnetism was Edmund Halley, who published, in 1683 and 1692, two works on the theory of geomagnetism, but needed to test his theory further. King William III put

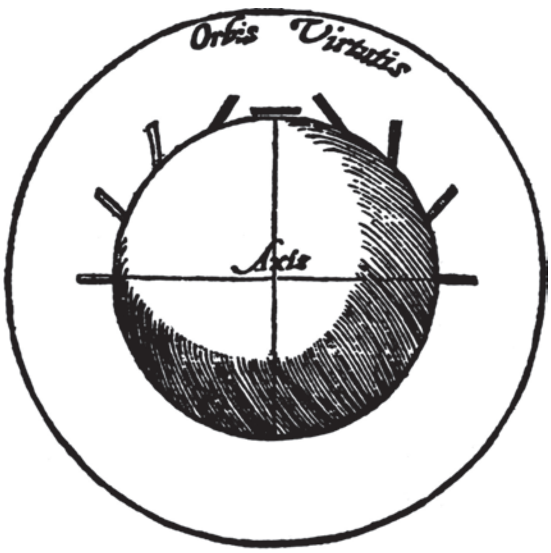

FIGURE 1.2. Illustration of the magnetic dipole character of Earth's main magnetic field, as shown in Gilbert's *De Magnete*.

at his disposal the ship *Paramour Pink*, on which Halley made two voyages: in October 1698 to the North Atlantic, and in September 1700 to the South Atlantic. Those voyages were the first purely scientific expeditions, and they returned measurements of great value, both for practical navigation and for the theory of navigation. Those investigations led to the publication of two geomagnetic charts published in 1701 and 1702, respectively: *New and Correct Chart showing the Variations of the Compass in the Western and Southern Oceans, as observed in year 1700 by his Majesty's Command by Edm. Halley* and *Sea Chart of the whole world, showing the Variations of the Compass.*

1.4 EMERGENCE OF A SCIENTIFIC DISCIPLINE

Despite the fact that the Sun is the most luminous object we can see, the solar half of solar-terrestrial physics awaited technological advance as surely as the study of geomagnetism awaited the development of the compass and its successor, the magnetometer. **Sunspots**, or magnetized cool spots in the solar photosphere, are generally too small to be resolved by the naked eye. Thus, the study of sunspots did not begin until the invention of the telescope. Galileo Galilei was one of the first to use this new invention in his study. Sunspot studies proceeded slowly, perhaps because very few sunspots occurred during the period known as the Maunder minimum, from about 1645 to 1700. After the Maunder minimum, the sunspot numbers were similar to those in more recent times, but, in the last decade of the 1700s, a remarkable change occurred. The solar cycle was much longer than usual, and the following two sunspot maxima were much smaller than before. This interval has been dubbed the Dalton minimum. At the turn of the millennium, a similar long solar cycle occurred and the solar activity plummeted, so it is possible that there are longer cycles of solar activity than that represented by the sunspots.

The now familiar 11 year periodicity in sunspot number, illustrated in Figure 1.3, was not discovered until 1851. The sunspot, or **solar cycle**, is discussed in greater depth in Chapter 4, which reviews our current understanding of the physics of the Sun, in which magnetism plays a significant role that is only gradually being understood. In fact, the deep solar minimum that occurred from 2006 to 2010 as cycle 23 ended was accompanied by weak magnetic fields on the Sun.

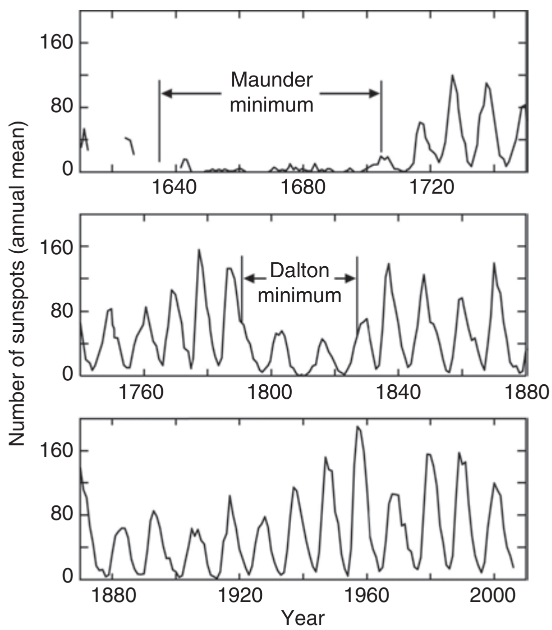

FIGURE 1.3. Sunspot cycle since AD 1610.

Perhaps the first discovery in the emerging discipline we now call solar-terrestrial physics was the observation in 1722 by George Graham, a famous London instrument maker, that the compass is always in motion. Graham's discovery was confirmed in 1740 by Anders Celsius in Uppsala, Sweden. His observations were continued by O. Hiorter to a total of over 20 000 observations made on more than 1000 different days. From those data, Hiorter discovered the diurnal variation of the geomagnetic field. Magnetic perturbations vary systematically with local time, which is determined by the longitudinal separation between the meridian of the observer and that containing the Sun, the noon meridian. These perturbations are due to the rotation of the observer under current systems flowing in the upper atmosphere that are fixed with respect to the Sun.

Even more importantly, on April 5, 1741, Hiorter discovered that geomagnetic and auroral activities were correlated. Simultaneous observations in London by Graham confirmed the occurrence of strong geomagnetic activity on that day. In 1770, J. C. Wilcke noted that auroral rays extend upward along the direction of the magnetic field. In the same year, Captain James Cook first reported the southern counterpart of the aurora borealis, the aurora australis, or "dawn of the south." Twenty years later, the English scientist Henry Cavendish used triangulation to estimate the height of auroras as between 80 and 115 kilometers. Earlier attempts at triangulation by Halley and Mairan had been much less accurate.

The great advance of the early nineteenth century was the development of a network to make frequent simultaneous observations with widely spaced **magnetometers**. C. F. Gauss was one of the leaders of that effort and one of the foremost pioneers in the mathematical analysis of the resulting measurements, which allowed contributions to the geomagnetic field from below the surface of the Earth to be separated from those contributions arising high in the atmosphere.

Meanwhile, Heinrich Schwabe, on the basis of his sunspot measurements taken between 1825 and 1850, deduced that the variation in the number of sunspots was periodic, with a period of about ten years. By 1839, magnetic observatories had spread to the British colonies. Edward Sabine was assigned to supervise four of those observatories (Toronto, St. Helena, Cape of Good Hope, and Hobarton) in 1851. Using these data, he was able to show that the intensity of geomagnetic disturbances varied in concert with the sunspot cycle. Chapter 9 discusses our modern understanding of these disturbances.

The next discovery linking the Sun and geomagnetic activity was Richard Carrington's sighting of a great white-light **flare** on September 1, 1859. Carrington, who was sketching sunspot groups at the time, was startled by the flare, and by the time he was able to summon someone to witness the event a minute later, he was dismayed to find that it had weakened greatly in intensity. Fortunately, it had been simultaneously noted by another observer some miles away. Furthermore, at the moment of the flare, measurements of the magnetic field at the Kew Observatory in London had been disturbed. Today, we realize that that disturbance of the magnetic field was caused by an increase in the electric currents flowing overhead in the Earth's ionosphere. Such currents flow in response to electric fields in the electrically conducting ionosphere. The radiation from extreme-ultraviolet rays and x-rays from the flare increased the ionization high in the atmosphere and hence the electrical conductivity, causing more current to flow in response to the unaltered electric field. Finally, 18 hours later, one of the strongest magnetic storms ever recorded occurred. Auroras were seen as far south as Puerto Rico. To have arrived that quickly, the disturbance would have had to travel from the Sun at over 2300 km s^{-1}. As discussed in Chapter 5, we know today that the Sun and the Earth are linked by the supersonic solar wind, but such a velocity is high even for the disturbed solar wind.

Total solar eclipses occur somewhere on the Earth almost annually, and these solar eclipses reveal the density structure of the solar corona. Now we know that the density structure can change rapidly and significantly in association with the solar events that lead to terrestrial storms, but total solar eclipses last only minutes, far too short to reveal these changes. Thus **coronal mass ejections**, such as the one that certainly led to Carrington's 1859 geomagnetic storm, remained undiscovered until the invention of the coronograph, which could make artificial solar eclipses for automatic cameras. Now the solar corona is regularly measured from high altitude and from space.

In 1861, shortly after Carrington's observations, Balfour Stewart noted the occurrence of pulsations in the Earth's magnetic field, with periods of minutes. Today, we know that the Earth's magnetic field pulsates at a wide variety of frequencies. These pulsations provide diagnostics of the state of the magnetosphere and the processes occurring therein. They are described in further detail in Chapter 10. The energy for these waves in the plasma can be exogenic, from the solar wind and its interaction with the magnetosphere, or endogenic, powered by the free energy in unstable plasmas produced in the dynamic magnetosphere. The exchange of energy between the electrons and ions in a plasma and waves that can transport the exchanged energy is a very active and important area of research in space plasma physics. This is described in more detail in Chapter 10.

The nineteenth century also brought another simple but important observation of auroras. Captain John Franklin, the ill-fated English Arctic explorer whose party perished in 1845 attempting to discover

Cambridge University Press
978-1-107-09882-4 - Space Physics
Christopher T. Russell, Janet G. Luhmann and Robert J. Strangeway
Excerpt
More information

the Northwest Passage (now open as a result of recent global warming), noted that auroral frequency did not increase all the way to the pole; this was according to observations made during his 1819–22 journeys. In 1860, Elias Loomis of Yale was one of the first to plot the zone of maximum auroral occurrence, which roughly corresponds to what we today refer to as the **auroral zone**. This zone is an oval band around the magnetic pole, roughly 20–25° from the pole. Precursors to our modern understanding of the auroras began to appear in the late nineteenth century. About 1878, H. Becquerel suggested that particles were shot off from the Sun and were guided by the Earth's magnetic field to the auroral zone. He believed that sunspots ejected protons. A similar theory was espoused by E. Goldstein.

In 1897, the great Norwegian physicist Kristian Birkeland made his first auroral expedition to northern Norway. However, it was not until after his third expedition in 1902–03, during which he obtained extensive data on the magnetic perturbations associated with auroras, that he concluded that large electric currents flowed along magnetic field lines during auroras. The invention of the vacuum tube led to the understanding that auroras were in some way similar to the cathode rays in those devices. Soon, Sir William Crookes demonstrated that cathode rays were bent by magnetic fields, and shortly thereafter J. J. Thomson showed that cathode rays consisted of the tiny, negatively charged particles we now call electrons. Birkeland adopted those ideas for his auroral theories and attempted to verify them with both field observations and laboratory experiments. Specifically, he conducted experiments with a magnetic dipole inside a model Earth, which he called a terrella. Figure 1.4 shows Birkeland in his laboratory beside his terrella experiment. Those experiments demonstrated that electrons incident on the terrella would produce patterns quite reminiscent of the auroral zone. He believed, as we do today, that those particles came from the Sun.

K. Birkeland's work inspired the Norwegian mathematician Carl Størmer, whose subsequent calculations of the motion of charged particles in a dipole magnetic field in turn supported Birkeland. Figure 1.5 shows Størmer and his assistant Bernt Johannes (not Kristian) Birkeland. As is evident from this

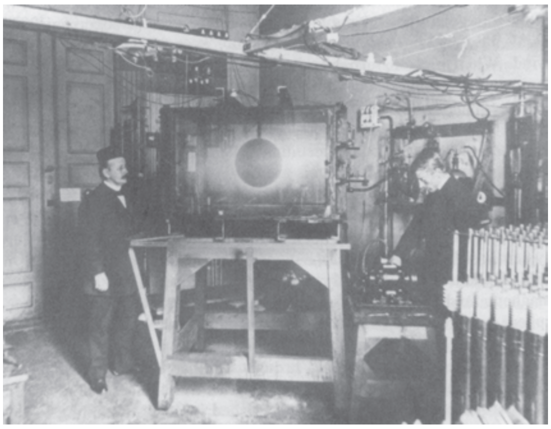

FIGURE 1.4. Kristian Birkeland (left) in his laboratory with his terrella and with his assistant, O. Devik (right), about 1909.

photograph, the advent of the camera was an important advance in the study of auroras. It was through measurements such as these that Størmer accurately determined the height of the auroras.

Figure 1.6 illustrates one of Størmer's charged-particle orbit calculations in a forbidden region to which charged particles from the Sun did not have direct access. In such a region, charged particles would spiral around the magnetic field and bounce back and forth along the field, reflected by the converging magnetic field geometry. Størmer's contributions became much more relevant and appreciated after the discovery of the Earth's radiation belts, whose particle motions resemble those of Figure 1.6. Kristian Birkeland's work was not appreciated until later, when spacecraft found field-aligned currents linked to the auroras. A more detailed discussion of the trajectories of charged particles in the Earth's magnetic field can be found in Chapter 10, and more about auroras can be found in Chapter 11.

This work on the solar-terrestrial connection proceeded, despite Lord Kelvin's 1882 argument that he had provided absolutely conclusive evidence against the supposition that terrestrial magnetic storms were due to magnetic action in the Sun or to any kind of dynamic action taking place within the Sun. Lord Kelvin also claimed "that the supposed connection between magnetic storms and sunspots is unreal, and the seeming agreement between periods has been a mere coincidence." More telling was the criticism of

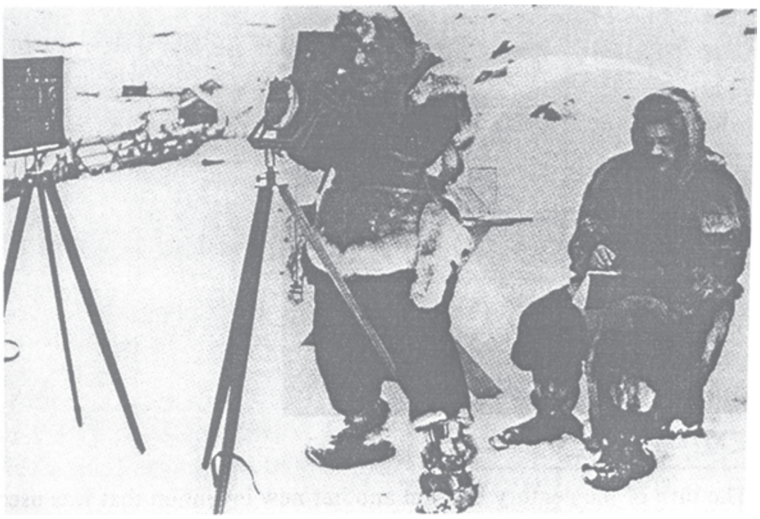

FIGURE 1.5. Auroral physicists C. F. Størmer, standing, and Bernt Johannes Birkeland, seated, in northern Norway, *c.* 1910.

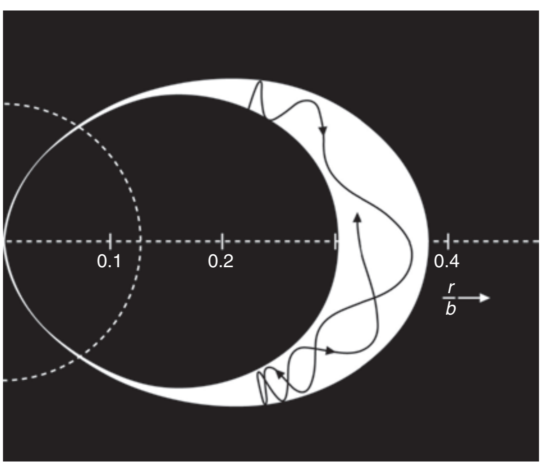

FIGURE 1.6. Trajectory of an energetic charged particle in the "forbidden" zone of a dipole magnetic field, as drawn by Størmer (1911). (After Rossi and Olbert, 1970.)

A. Schuster that a beam of electrons from the Sun could not hold together against their mutual electrostatic repulsion.

1.5 THE IONOSPHERE AND MAGNETOSPHERE REVEALED

The electrically conducting region above roughly 100 km altitude that we now call the **ionosphere** may rightly be claimed to have been discovered by Balfour Stewart. In his 1882 *Encyclopedia Britannica* article, entitled "Terrestrial magnetism," he concluded that the upper atmosphere was the most probable location of the electric currents that produce the solar-controlled variation in the magnetic field measured at the surface of the Earth. He noted that "we know from our study of aurora that there are such currents in these regions – continuous near the poles and occasional in lower latitudes."

He proposed that the primary causes of the daily variations in the intensity of the surface magnetic field were "convective currents established by the Sun's heating influence in the upper regions of the atmosphere." These currents "are to be regarded as conductors moving across lines of magnetic force and are thus the vehicle of electric currents which act upon the magnetic field." Those statements are very close to modern atmospheric-dynamo theory. However, it was left to A. Schuster to put the dynamo theory into quantitative form.

The turn of the twentieth century brought another new invention that was used to probe the solar-terrestrial environment: the radio transmitter and receiver. In 1902, A. E. Kennelly and O. Heaviside independently postulated the existence of a highly electrically conducting ionosphere to explain G. Marconi's transatlantic radio transmissions. Verification of the existence of the ionosphere did not come until much later, in 1925, when E. V. Appleton and M. A. F. Barnett in the United Kingdom, and shortly thereafter G. Breit and M. A. Tuve in the United States, established the existence and altitude of the Kennelly–Heaviside layer, as it was known then. The only contemporary usage of the term "Heaviside layer" occurs in the musical *Cats*.

The original method of Breit and Tuve, using short pulses of radio energy at vertical incidence and timing the arrival of a reflected signal in order to infer the altitude of the electrically reflecting layer, is still used today for sounding the ionosphere. In drawing diagrams of the electromagnetic waves reflected by the ionosphere, Appleton used the letter E for the electric vector of the downcoming wave. When he found reflections from a higher layer, he used the letter F for the electric vector of those reflected waves, and when he occasionally got reflections from a lower layer, he naturally used the letter D. When it came time to name these layers, he chose the same letters, leaving the letters A, B, and C for possible later discoveries that never came. So, today, the ionospheric layers are referred to as the D, E, and F layers, as illustrated in Figure 1.7. We now know that all planets with atmospheres have electrically conducting ionospheres. Chapter 2 discusses how these are formed.

At about that same time, progress was being made in understanding the auroral glows. Spectroscopy, together with photography, permitted first the determination of the wavelength and then the identity of the excited molecules that were radiating. There were initial successes, beginning with Lars Vegard's work in Norway, relating auroral emissions to emission bands from known atmospheric gases such as nitrogen. However, identification of the yellow-green line at 557.7 nm was elusive. Finally, H. Babcock's precise measurements in 1923 allowed John McLennan to identify this line as a metastable transition of atomic oxygen. At atmospheric pressures close to that at the surface of the Earth, collisions between molecules de-excite the molecules before they have a chance to radiate if they happen to become excited into one of the metastable states. However, at the altitude of the auroras, collisions are so rare that the time between collisions is longer than the lifetimes of the metastable states, and the excitation energy of the state can be released by radiation. A similar line in the auroral spectrum is the 630.0 nm red line of atomic oxygen. This metastable transition has a lifetime of 110 s and can radiate only above some 250 km. Those discoveries led to the realization that the varied colors of the auroras were related to height. In low-altitude auroras, below 100 km, where collisions quench even the oxygen green line, the blue and red nitrogen bands predominate. From 100 to 250 km, the oxygen green line is strongest. Above 250 km, the oxygen red line is most important.

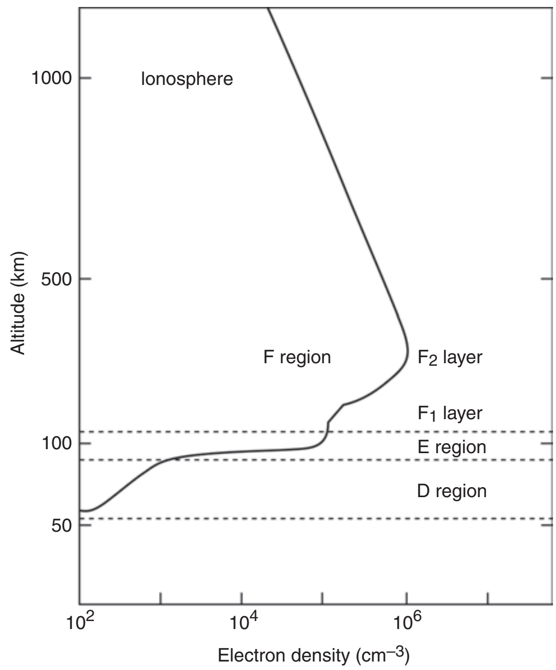

FIGURE 1.7. Electron density of the Earth's ionosphere as a function of altitude.

Although most auroral forms are associated with electrons, some auroras are due to precipitating protons. The first observations of the proton aurora were made in 1939. Measurements of the Doppler shifts of the proton emissions permitted estimates of the energy of the precipitating particles from the ground. Chapter 11 contains a more detailed discussion of auroras and the auroral ionosphere.

With the concept of the ionosphere firmly established, scientists began to wonder about the upper extension of the ionosphere that is linked magnetically to the Earth; today, this is called the magnetosphere. In 1918, Sydney Chapman postulated a singly charged beam from the Sun as the cause of worldwide magnetic disturbances, a revival of an old idea that had previously been criticized by Schuster. Chapman was soon challenged by Frederick Lindemann, who pointed out that mutual electrostatic repulsion would destroy such a stream. Lindemann instead suggested that the stream of charged particles contained particles of both signs in equal numbers. We now call the material in such a stream a "**plasma.**" That proposal was a breakthrough, and it permitted Chapman and his co-workers, in a series of papers beginning in 1930, to lay the foundations for our modern understanding of the interaction of the solar wind with the magnetosphere. Today we understand plasmas sufficiently to use plasma beams to both orient and propel our spacecraft, but it was not until the 1960s that the acceleration of the solar plasma was understood.

In the rarefied conditions of outer space, where collisions between particles are infrequent, the ion–electron gas, or plasma, is highly electrically conducting. Thus, Chapman and Ferraro proposed that, as the plasma from the Sun approached, the Earth would effectively see a mirror magnetic dipole moment advancing, as illustrated in Figure 1.8. The net result of that advancing mirror field would be to compress the terrestrial field. Eventually, as sketched in Figure 1.9, the plasma would surround the Earth on all sides, and a cavity would be carved out of the solar plasma by the terrestrial magnetic field. That is very similar to our modern concept of the geomagnetic cavity, the formation of which is discussed in greater detail in Chapter 7.

After the compression of the **magnetosphere**, which is detected by ground-based magnetometers as a sharp increase in the magnetic field, the magnetosphere becomes inflated. Chapman and Ferraro correctly interpreted this subsequent decrease in the magnetic field at the surface of the Earth as the appearance of energetic plasma deep inside the magnetosphere, forming a ring of current in the near-equatorial regions. The development of this ring current in what we now call a geomagnetic storm is discussed at greater length in Chapter 9.

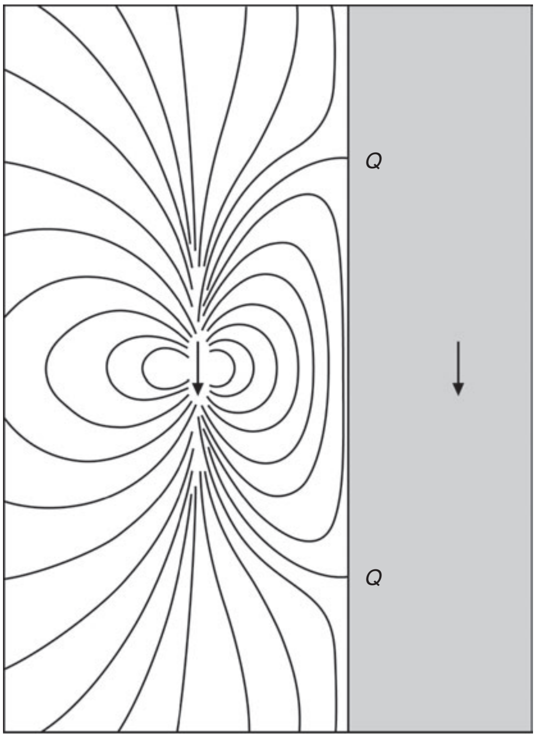

FIGURE 1.8. Compression of a dipole field by an advancing infinite, superconducting slab. The magnetic field is due to the original dipole plus an image dipole an equal distance behind the front, as shown by the right-hand arrow. (After Chapman and Bartels, 1940.)

At the same time as the ionosphere was being discovered by virtue of its effects on man-made radio signals, natural radio emissions were also being explored, and the magneto-ionic theory, developed for the man-made signals, was being applied to those natural emissions. The first report of electromagnetic signals in the audio-frequency range was an observation of what have become known as "**whistlers,**" detected using a 22 km telephone line in Austria in 1886. Whistlers are short bursts of audio-frequency radio noise of continuously decreasing pitch. In 1894, British telephone operators heard "tweeks," possibly whistlers generated by lightning, and a "dawn chorus" generated deep in the magnetosphere during a display of the aurora borealis. Little work was done on those observations because of the lack of suitable analysis equipment at the time. During World War I,

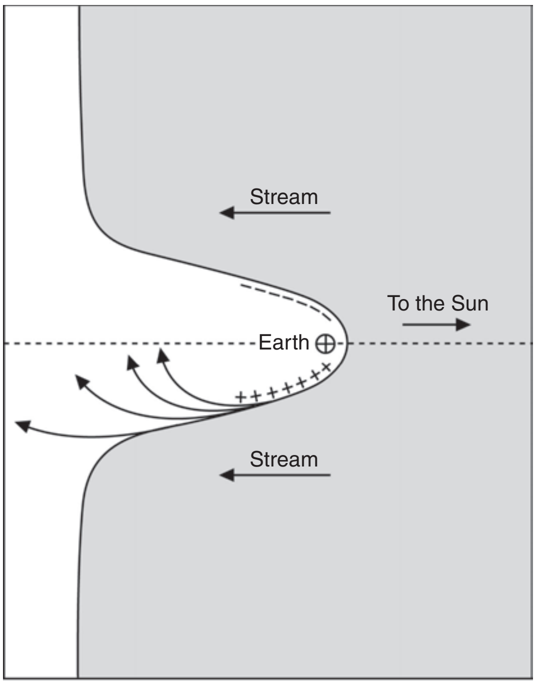

FIGURE 1.9. Expected evolution of the front of superconducting plasma as it passes the Earth. This model was proposed by Chapman and Ferraro in the 1930s to explain the phenomena of the geomagnetic storm. (After Chapman and Bartels, 1940.)

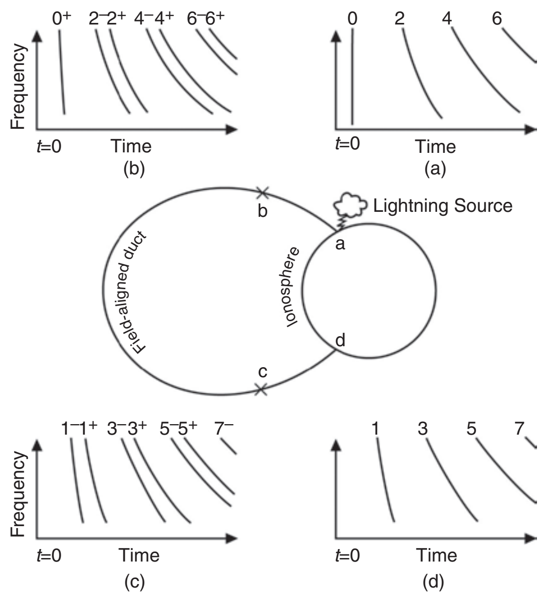

FIGURE 1.10. Dispersion of whistler-mode waves generated by lightning, as seen at four different locations. The different velocities of propagation as functions of wave frequency (dispersion) cause the wave arrival to be delayed by a different amount at each frequency. The delay depends on the distance traveled and the properties of the plasma traversed by the wave. (From Russell, 1972.)

equipment installed to eavesdrop on enemy telephone conversations picked up whistling sounds. Soldiers on the front would say, "You could hear the grenades fly." H. Barkhausen reported on those observations in 1919, and suggested that they were correlated with meteorological influences. However, he could not duplicate the phenomenon in laboratory experiments.

In 1925, T. L. Eckersley also described that phenomenon but incorrectly ascribed it to the dispersion of an electrical impulse in a medium loaded with free ions, causing signals at different frequencies to travel at different speeds. Eventually, in 1935, after much work and several incorrect explanations, Eckersley concluded that the distinctive swooping sound of whistlers was due to electron-caused dispersion of a burst of electromagnetic noise traveling through the ionosphere.

Very little work was done on whistlers until the early 1950s, at which time L. R. O. Storey, with a homemade spectrum analyzer, conducted a thorough study of whistlers. He found that whistlers are caused by lightning flashes, whose electromagnetic energy then echoes back and forth along field lines in the upper ionosphere, as illustrated in Figure 1.10. A major implication of those findings was that the electron density in the outer ionosphere, which is now called the plasmasphere, was unexpectedly high. Storey also found other types of audio-frequency, or very low frequency (VLF), emissions that are not associated with lightning and are now known to be generated within the magnetospheric plasma. Chapters 10 and 13 discuss the generation and propagation of these waves.

Because of atmospheric drag, it is difficult to study the upper atmosphere and ionosphere from space. Table 1.1 lists several missions that have attempted to explore part of this low-altitude environment. The Alouette 1 and 2 spacecraft used radio sounding to study the topside ionosphere. ISIS 1, 2