The High-Latitude Ionosphere and its Effects on Radio Propagation

R. D. Hunsucker, J. K. Hargreaves

Book DOI: http://dx.doi.org/10.1017/CBO9780511535758

Online ISBN: 9780511535758

Hardback ISBN: 9780521330831

Paperback ISBN: 9780521041362

Chapter

Chapter 2 - Geophysical phenomena influencing the high-latitude ionosp

here pp. 61-112

Chapter DOI: http://dx.doi.org/10.1017/CBO9780511535758.004

Cambridge University Press

Chapter 2

Geophysical phenomena influencing the high-latitude ionosphere

2.1 Introduction

Whereas the mid-latitude ionosphere is dominated by solar radiation and the chemistry of the upper atmosphere, modified by dynamic effects, the high-latitude ionosphere is, in addition, strongly affected by the nature of the geophysical environment and by various processes occurring within it. In particular, the form of the geomagnetic field connects the polar upper atmosphere to the magnetosphere. Thereby the polar ionosphere becomes accessible to particles that have been energized within the magnetosphere or have come from the Sun; these provide another source of ionization. It is also affected by the dynamics of the magnetosphere and is thus subject to electric fields and currents generated by motions at high levels. At the highest latitudes the ionosphere is connected, via the field-lines, to the outer magnetosphere, giving it a ready response to variations in the flow of the solar wind.

The present chapter therefore summarizes the basic properties and behavior of the magnetosphere, which we must appreciate in order to understand the behavior of the ionosphere poleward of 60° latitude.

2.2 The magnetosphere

2.2.1 The geomagnetic field

To a first approximation the geomagnetic field at and close to the planet's surface can be represented as a dipole field. The poles of the dipole are at geographic

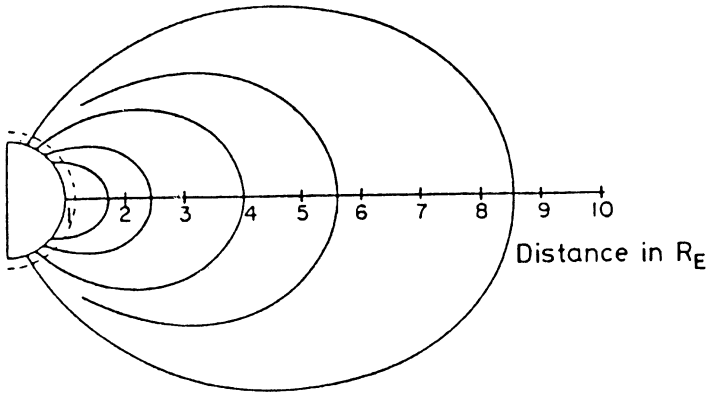


Figure 2.1. Dipolar field-lines. (D. L. Carpenter and R. L. Smith, *Rev. Geophys.* **2**, 415, 1964, copyright by the American Geophysical Union.)

latitudes and longitudes 79° N, 70° W, and 79° S, 70° E. The magnetic flux density is given by

$$B(r, \lambda) = \frac{M}{r^3}(1 + 3\sin^2\lambda)^{1/2}, \tag{2.1}$$

where M is the dipole moment, r the geocentric radial distance, and λ the magnetic latitude. This is accurate to within about 30% at points within two or three Earth-radii of the surface. Although it is not very accurate, the dipole form is useful for making approximate calculations.

Figure 2.1 shows the lines of force, generally called *field-lines*, in a dipole field. Each line is the locus of the force on a single north pole and is represented by a simple equation,

$$r = r_0 \cos^2\lambda. \tag{2.2}$$

If $\lambda = 0$, $r = r_0$; r_0 is thus the radial distance to the field-line in the plane normal to the axis of the dipole. There is a different value of r_0 for each line of Figure 2.1, but, since the field is three-dimensional, each r_0 actually describes a shell. The other coordinate is provided by *magnetic longitude*.

In the magnetosphere it is convenient to use the radius of the Earth, R_E, as the unit of distance. Then, putting $r/R_E = R$,

$$B(R, \lambda) = \frac{0.31}{R^3}(1 + 3\sin^2\lambda)^{1/2}\ \text{G}. \tag{2.3}$$

0.31 G ($= 3.1 \times 10^{-5}$ Wb m^{-2}) is the flux density at the magnetic equator on the Earth's surface. In these terms, the field-line equation becomes

$$R = R_0 \cos^2\lambda, \tag{2.4}$$

where both R and R_0 are measured in Earth-radii. The latitude where the field-line intersects the Earth's surface is given by

$$\cos \lambda_E = R_0^{-1/2}. \tag{2.5}$$

The dipole form is convenient for its mathematical simplicity, but for many purposes it is not sufficently accurate. A closer approximation is the displaced-dipole model, in which the dipole is displaced by 400 km from the center of the Earth. However, for accurate work (not too far above the surface) it is usual to derive the field from the magnetic potential expressed as a series of spherical harmonics – the flux density being the gradient of the magnetic potential. The dipole form corresponds to the first term of the expansion.

The coefficients are derived by fitting the expression to measurements of the magnetic elements on the global scale, using magnetometers both on the ground and on satellites. Because the geomagnetic field changes with time – the secular variation – a fresh set of coefficients, relating to a specific epoch, is published from time to time. Such representations are accurate to within about 0.5% at and near the surface. The terms of higher order become less important at greater distances and the field tends to become more dipolar. However, beyond three or four Earth-radii the distortion due to the solar wind has to be taken increasingly into account.

The pressure of the solar wind confines the geomagnetic field on the sunward side and forms the *geomagnetic cavity*.

2.2.2 The solar wind

The solar wind was first observed directly by space probes in the early 1960s, though its existence had previously been proposed in theoretical work and some of its properties had been deduced from studies of comets.

There have been many observations of the solar wind since that time. It is basically an outflow driven by the continual expansion of the solar corona and it is therefore composed of solar material. Most of the ions are protons (H^+) but there is also an α-particle (He^{2+}) component typically amounting to 5% though exceptionally up to 20%. Still heavier atoms amount to perhaps 0.5% in total, though, in contrast to the light ions, these are not fully ionized. The concentration of positive ions varies between 3 and 10 cm^{-3} (3×10^6 to 10^7 m^{-3}), the most typical value being 5 cm^{-3}, and there is a similar number of electrons for bulk neutrality. The mean mass of solar-wind particles is therefore about half that of the proton, about 10^{-27} kg. There are fluctuations as large as by a factor of ten over times of minutes and hours, implying irregularities within the solar wind over distances of 10^5 km and more.

At the distance of the Earth's orbit the speed of the solar wind is usually between 200 and 700 or 800 km s^{-1} (Figure 2.2), on which is superimposed a random component of temperature 10^5 K. The solar wind is not very hot by solar

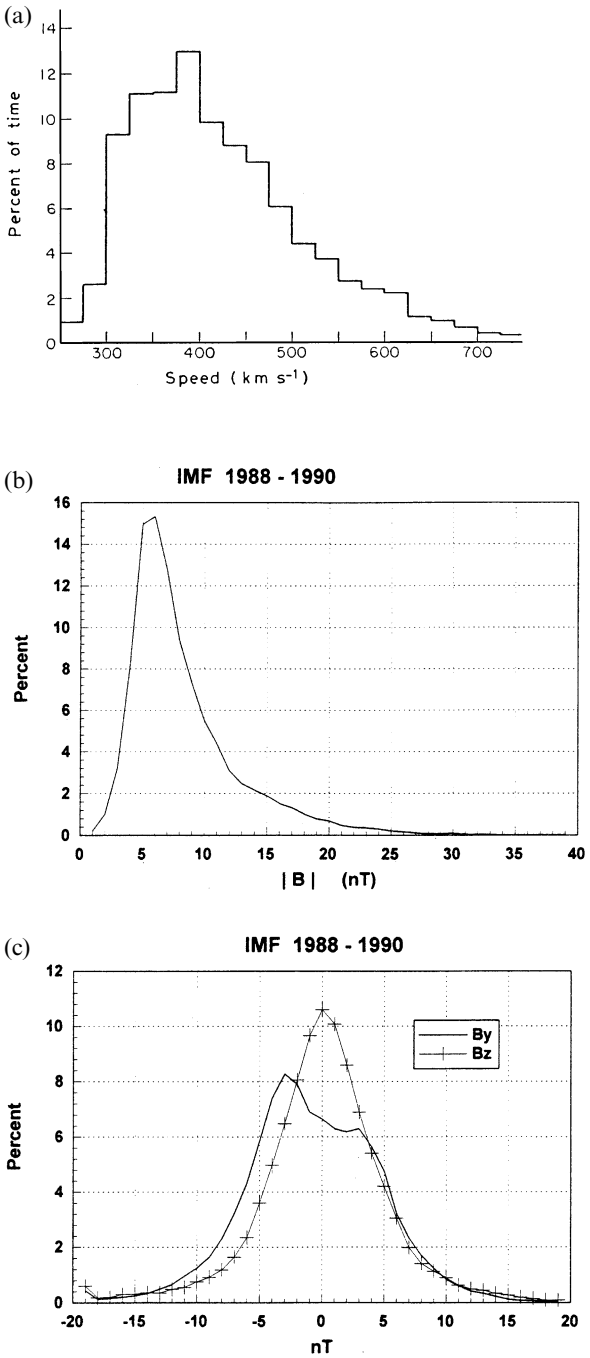


Figure 2.2. (a) The speed of the solar wind: a histogram of measurements between 1962 and 1970. (J. T. Gosling, in *Solar Activity Observations and Predictions* (eds. McIntosh and Dryer), by kind permission of The MIT Press, 1972.) In (b) (c) are shown the distributions of the magnitude and components of the interplanetary magnetic field, 1988–1990. B_y is east–west and B_z is north–south. (F. J. Rich and M. Hairston. *J. Geophys. Res.* **99**, 3827, 1994, copyright by the American Geophysical Union.)

standards, the energy being more directed than random. It carries an energy flux of about 10^{-4} W m^{-2}, which is approximately a tenth of that in the EUV region of the solar spectrum.

The solar wind is the principal medium by which the activity of the Sun is communicated to the vicinity of the Earth, and it is extremely important in solar–terrestrial relations and in the behavior of the high-latitude ionosphere.

The interaction depends on a weak magnetic field, the *interplanetary magnetic field (IMF)*, which is carried along by the plasma. This field amounts only to a few nanoteslas (a few γ) and it is "frozen in" to the plasma because of the large electrical conductivity. The magnitude of the IMF varies slightly with the sunspot cycle (Figure 1.17). The kinetic energy of the solar-wind particles exceeds the energy density of the magnetic field by a factor of about eight, and therefore the motion of the total magnetoplasma is governed by the motion of the particles rather than by the magnetic field.

Although the solar wind flows out almost radially from the Sun, the solar rotation gives the magnetic field a spiral form, as in Figure 2.3. This is sometimes known as the *garden-hose effect* since it may be simulated by turning round while watering the garden and noting that the jet of water follows a spiral path although the trajectories of individual drops are radial. It so happens that, at the orbit of Earth, the IMF field-lines run at about 45° to the radial direction: the radial and the east–west components of the IMF are therefore about equal in magnitude.

One of the most remarkable of the early results, and a fact of great significance, is that distinct sectors may be recognized within the solar wind, the field being inward and outward in alternate sectors. Figure 2.3 shows some of the original measurements, in which four sectors – two inward and two outward – were present. However, this is not always the case because the sector structure evolves with time. Sometimes there are only two sectors, and sometimes the sectors are not all of the same width.

The proton density can vary by more than a factor of ten and the speed of the solar wind by a factor of two during one solar rotation as the sectors go by, with a degree of anticorrelation.

At first sight the form of the IMF appears anomalous. Although there may be a north–south component, it is equally likely to be northward or southward; thus it seems that a spiral in the ecliptic plane is indeed the basic form of the IMF. How is this to be reconciled with an origin in the solar magnetic field which we expect to be essentially dipolar? The problem is that the early observations were confined to the ecliptic plane and there is still not much knowledge of its form at higher solar latitudes. It is now thought that there is a current sheet in or near the equatorial plane that effectively divides the outward field (above the plane) from the inward field (below it) as in Figure 2.4. If the solar magnetic dipole is tilted from the rotation axis, the current sheet will be tilted from the ecliptic plane and a spacecraft near the Earth will observe a two-sector structure as the Sun rotates. When more than two sectors are seen, it is thought that the current sheet has developed

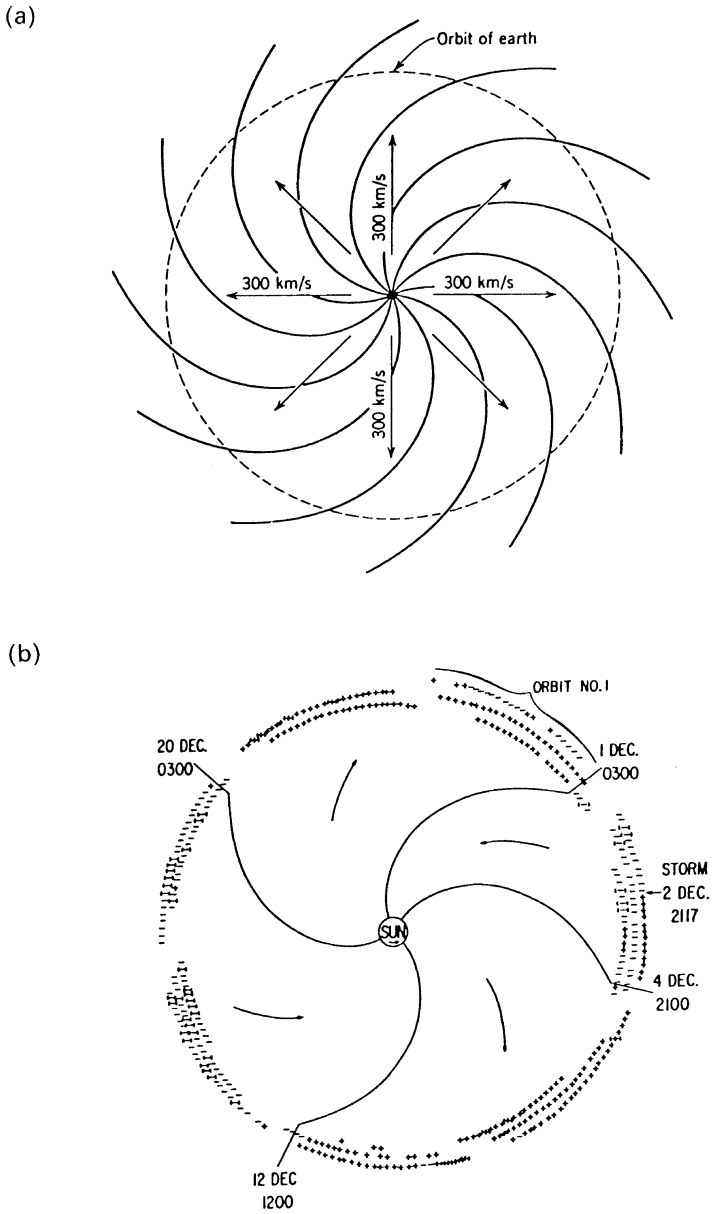


Figure 2.3. (a) The form of the interplanetary magnetic field (IMF) in the solar equatorial plane, corresponding to a solar-wind speed of 300 km s^{-1}. (T. E. Holzer, *Solar System Plasma Physics*, *Vol. I*, North-Holland, 1979, p. 103, Elsevier Science Publishers.) (b) The sector structure of the solar wind in late 1963, showing inward (−) and outward (+) IMF. (J. M. Wilcox and N. F. Ness, *J. Geophys. Res*. **70**, 5793, 1965, copyright by the American Geophysical Union.)

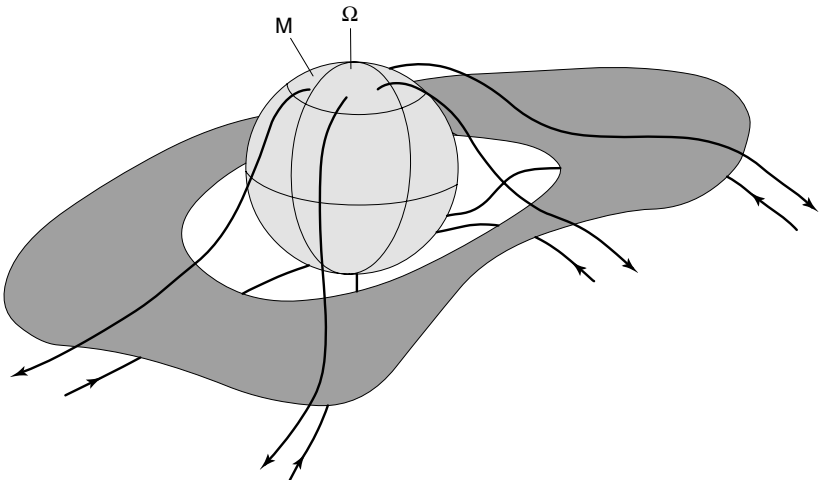


Figure 2.4. The ballerina model of the current sheet in the solar wind. M is the axis of the current sheet and Ω is the Sun's rotation axis. (E. J. Smith *et al.*, *J. Geophys. Res.*, **83**, 717, 1978, copyright by the American Geophysical Union.)

undulations as in the skirt of a pirouetting ballerina; hence the concept of Figure 2.4 is often known as the *ballerina model*. Spacecraft venturing out of the ecliptic plane have observed that the sector structure disappears – which is consistent with the ballerina model.

A link between the solar wind and a particular feature of the corona was discovered by the Skylab missions between May 1973 and February 1974. A so-called *coronal hole* emits less light at all wavelengths than do adjoining regions, but it is most marked in an X-ray photograph, on which it appears as a black area. Coronal holes are regions with abnormally low density where the magnetic field has a single polarity – all inward or all outward. This is an open magnetic field that goes out into interplanetary space rather than returning to the Sun. The hole is the source of fast solar-wind streams in which the speed exceeds 700 km s^{-1}. The speed is greater from a larger hole. Less than 20% of the solar surface is composed of coronal holes, and they are more numerous during the declining phase of the sunspot cycle.

The fast streams interact with the slower solar wind as in Figure 2.5(a), compressing the magnetic field and the plasma ahead and sometimes, though not always, creating a shock front. The compressed plasma is heated, and a rarefaction follows. Within the stream the magnetic field maintains the same polarity (inward or outward) and is the same as in the corresponding coronal hole. The fast streams from coronal holes co-rotate with the Sun and can persist for several rotations. They are the probable cause of recurring geomagnetic storms (Section 2.5.4).

Intermittant perturbations of the solar wind can be caused by specific solar events, particularly the *coronal mass ejection (CME)*. This is not the same as a

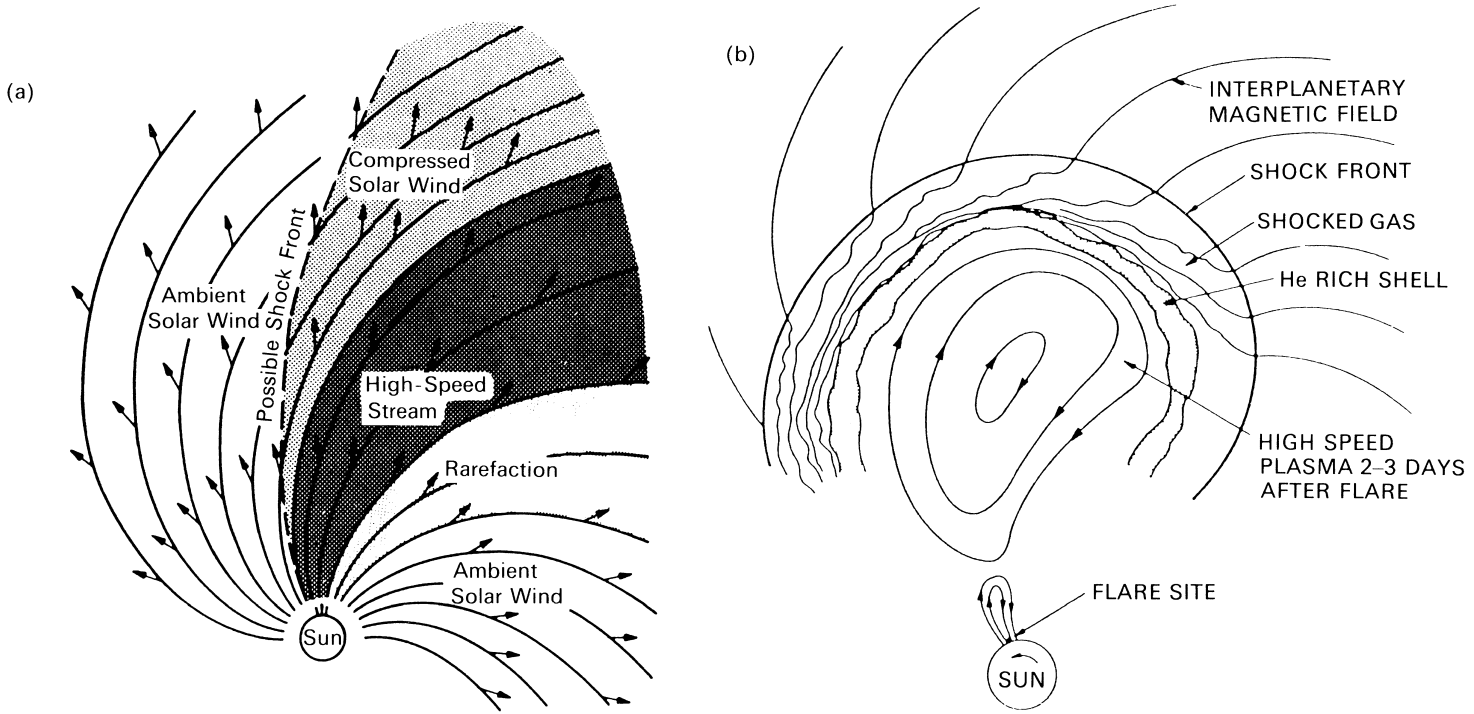


Figure 2.5. (a) Interaction between a localized stream of high-speed plasma and the slower, ambient solar wind. (T. E. Holzer, in *Solar System Plasma Physics, Vol I*, North-Holland, 1979, p. 103, Elsevier Science Publishers.) (b) High-speed plasma from a solar flare driving an interplanetary shock. The ejected plasma contains an ordered magnetic field, but between the shock and the ejecta the field is turbulent. (Reprinted from L. F. Burlaga, *Adv. Space Res*. **2**, 51, copyright 1982, with permission from Elsevier Sceince.)

solar flare, though in some instances a flare occurs at about the same time or shortly afterwards. The CME travels away from the Sun at a speed that may be less than 50 km s^{-1} or greater than 1200 km s^{-1}, and the speedier examples produce a shock front in the solar wind. The typical structure of such a disturbance is illustrated by Figure 2.5(b), (except that "flare" should be replaced by "CME"). The IMF is compressed by the shock, and a turbulent region is formed between the shock and the ejected matter. Within the CME the magnetic field is strong and well ordered, possibly as a closed loop. These magnetic structures, sometimes called *magnetic clouds*, are about 0.25 AU across at the orbit of Earth.

The form of the cavity formed by the interaction of the solar wind and the geomagnetic field is illustrated in Figure 2.6. Because it has very high electrical conductivity, the solar wind is not able to penetrate the geomagnetic field but is swept around it. Pressure is exerted against the magnetic field, which is distorted and confined within a large but nevertheless limited region around the Earth. This kind of behavior was foreseen by Chapman and Ferraro as long ago as 1930 in their pioneering study of the cause of magnetic storms (Section 2.5.2). (In modern terms the solar wind is said to be "frozen out" of the geomagnetic field.)

The magnetosphere has a complex structure. In the rest of this section we will describe some of its main features: the magnetopause, the magnetosheath and the shock, the polar cusps, and the magnetotail. To begin with they will be treated as though they were essentially static. Dynamic aspects will be introduced in Section 2.4.

2.2.3 The magnetopause

To a first approximation the form of the boundary between the geomagnetic field and the incident solar wind can be deduced by considering the pressure balance across the boundary. We assume that, when the system is in equilibrium, the pressure of the solar wind outside is at every point of the surface equal to that of the magnetic field inside. If the solar wind contains N particles m^{-3}, each of mass m kg, traveling at velocity v m s^{-1} and striking the surface at angle ψ from the normal, then it can be shown that the total rate of change of momentum due to the flux of solar wind particles is $2Nmv^2\cos^2\psi$ N m^{-2}. This has to be equated to the magnetic pressure $B^2/(2\mu_0)$. All species within the solar wind contribute but the protons have greatest effect.

A simple calculation along these lines readily gives a realistic distance for the position of the boundary (approximately $10R_E$) along the Earth–Sun line, and allows one to estimate how it varies if the solar wind changes. We assume that $\psi = 0$, and the magnetic flux density B varies as (distance)$^{-3}$. Then the distance to the magnetopause is

$$l_m = \left(\frac{B_E^2}{4\mu_0 N m v^2} \right)^{1/6}, \tag{2.6}$$

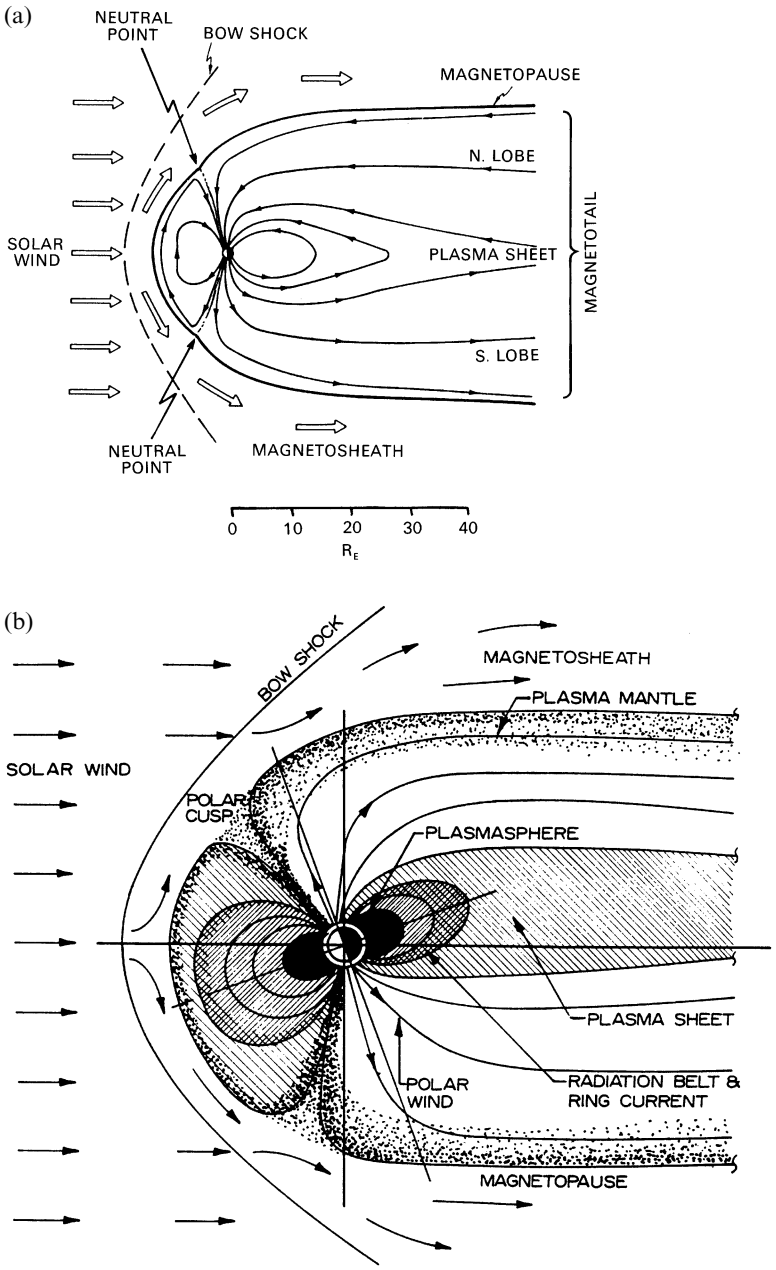


Figure 2.6. Two sketches of the geomagnetic cavity in north–south cross-section. (a) Showing the external flow of solar-wind plasma and the principal features of the distorted geomagnetic field. Note the tapering of the inner field-lines, suggesting that there is a neutral line further down the tail. (Adapted from V. M. Vasyliunas, in *Solar–Terrestrial Physics*, Reidel, 1983, p. 243, with kind permission from Kluwer Academic Publishers.) (b) Showing the main plasma regions in relation to the magnetic structure. Note the northward displacement of the plasma sheet during summer in the northern hemisphere. (W. J. Raitt and R. W. Schunk, in *Energetic Ion Composition in the Earth's Magnetosphere* (ed. R. G. Johnson), Terra Scientific Publishing Co., Tokyo, 1983, p. 99. After Bahnsen 1978).)

where B_E is the geomagnetic flux density at the Earth's surface at the magnetic equator.

A full computation is more complicated since the orientation of the boundary at each point is not known at the outset and an iteration proceedure is required. There is also a degree of coupling between the interplanetary and geomagnetic fields, which affects the location of the boundary. Spacecraft find the magnetopause about $0.5R_E$ closer to the Earth when the IMF has a southward component rather than a northward one. According to Petrinec and Russell (1993) the boundary moves closer by one R_E for every 7.4 nT of southward IMF component, but the distance is not much affected by varying amounts of northward component.

Another approximate method, which is fairly successful over the sunward side of the magnetopause only, is the *image-dipole* method, which replaces the dynamic pressure of the solar wind by the magnetic pressure of an image dipole of moment M_I placed parallel to and at distance d from the Earth's dipole, M_E. The fields due to these two dipoles are added and the distorted field-lines associated with M_E – the two fields do not interconnect – are taken to represent the geomagnetic field within the magnetopause. Those associated with the image have no physical significance. A satisfactory model is given by $M_I = 28M_E$; $d = 40R_E$. This method is not valid down the sides of the magnetosphere and in the anti-sunward direction.

The resulting boundary, the *magnetopause*, is indicated in Figure 2.6. The geomagnetic field is severely distorted within the magnetosphere. Note in particular the following points.

(a). Field-lines originating at low latitude form closed loops between northern and southern hemispheres, though there can be some distortion from the dipole form.

(b). Lines emerging from the polar regions are swept back, away from the Sun; in a dipole field some of these would have connected on the day side.

(c). Intermediate between these regions are two lines, one in each hemisphere, that go out and meet the magnetopause on the day side, though in fact their flux density falls to zero as they reach it; here, neutral points are formed.

The magnetopause has a finite thickness, though it is thin (approximately 1 km thick) in comparison with the size of the magnetosphere. Figure 2.7 (in Section 2.3.1) gives some idea of the form of the magnetosphere in three dimensions.

2.2.4 The magnetosheath and the shock

A shock front is formed in the solar wind two or three R_E upstream of the magnetopause (Figure 2.6). The region between the shock front and the magnetopause is the *magnetosheath*, and here the plasma, composed mainly of solar material but in other respects not typical of the solar wind, is turbulent.

Tenuous though it may be by any ordinary standard, the magnetosphere is a

relatively solid object in comparison with the solar plasma. Furthermore, the solar wind is "supersonic" at the orbit of Earth, meaning that its velocity exceeds that of any waves that can propagate within it. In the solar wind the speed of hydromagnetic waves, that is, the Alfvén speed, given by

$$v_A = \frac{B}{(\mu_0 \rho)^{1/2}}, \tag{2.7}$$

where B is the magnetic flux density and ρ the particle density (in kg m^{-3}), is about 50 km s^{-1}. For a solar wind speed of 400 km s^{-1}, therefore, the *Alfvén Mach number* is 8. We therefore have the conditions for a shock front, a discontinuity created when information about an approaching obstruction is not transmitted ahead into the medium.

The existence and location of the shock were predicted from theory in the early 1960s and subsequently verified by observation. On crossing the shock, solar-wind plasma is slowed down to about 250 km s^{-1} and the corresponding loss of directed kinetic energy is dissipated as thermal energy, increasing the temperature to 5×10^6 K. Magnetosheath plasma is therefore slower than the solar wind proper but 5–10 times hotter.

2.2.5 The polar cusps

The simple models of the magnetosphere predict two neutral points on the magnetopause where the total field is zero. These points connect along field-lines to places on the Earth's surface near $\pm 78°$ magnetic latitude. These are in fact the only points on the Earth's surface which connect directly to the magnetopause, and all the field from the magnetopause converges into those two points. They are therefore regions of great interest where solar-wind particles (from the magnetosheath) can enter the magnetosphere without having to cross field-lines.

Measurements reveal regions that are more extended than points, and they are now called the *polar cusps* or *clefts*. Particles with energies typical of those in the sheath are observed over some 5° of latitude around 77°, and over 8 h of local time around noon. The cusps are funnels of weak magnetic field filled with magnetosheath plasma, and they have significant effects on the high-latitude ionosphere. The ionospheric effects of particle entry provide "signatures" of the cusp, indicating its location – see Section 5.2.2 and Figure 5.7.

2.2.6 The magnetotail

In the anti-sunward direction the magnetosphere is extended into a long tail, the *magnetotail*. As is shown by spacecraft magnetometers, the geomagnetic field beyond about $10R_E$ on the night side of the Earth tends to run in the Sun–Earth direction, and there is a central plane within which the field reverses direction. This is the *neutral sheet*. The field points towards the Earth in the northern *lobe*,

and away from the Earth in the southern. The tail is roughly circular, some $30R_E$ (2×10^5 km) across, and of uncertain length, though it has been detected downwind beyond 10^7 km. Its significance for the high-latitude ionosphere is that it maps into the polar caps at its earthward end, and thus the polar ionosphere can be affected directly by events in the tail.

The basic form of the magnetotail in the plane containing the magnetic poles is shown in Figure 2.6. The flux density is about 20γ (= 20 nT) in the tail lobes, but the field is much weaker in the neutral sheet where the reversal occurs. In this region the magnetic pressure of the tail lobes ($B_T^2/(2\mu_0)$) is more or less balanced by an enhancement of the plasma density, the *plasma sheet* (to be considered further in Section 2.3.3). However, in fact the tail, like the whole magnetosphere, is dynamic and it forms an essential part of the magnetospheric circulation, to be considered in Section 2.4.

2.3 Particles in the magnetosphere

2.3.1 Principal particle populations

The geomagnetic field holds within it several distinct populations of charged particles.

(a). Deep within the magnetosphere (in the region often known as the inner magnetosphere) is the *plasmasphere*, closely linked to the mid-latitude ionosphere and comprising electrons, protons, and some heavy ions, all having energies in the thermal range.

(b). Also trapped on closed field-lines are the energetic particles generally known as the *Van Allen particles* after their discoverer. Apart from cosmic rays and solar protons, which are merely passing through, the Van Allen particles are the most energetic particles in the magnetosphere and they make some contribution to the ionization of the upper atmosphere when they are precipitated out of the trapping region.

(c). The *plasma sheet* is associated with the magnetotail, essentially with the central region where the magnetic field reverses direction. Plasma-sheet particles are energized within the magnetotail and they are important in auroral activity and the behavior of the high-latitude ionosphere. Their energy is intermediate between those of the plasmasphere and the Van Allen belt. The inner edge of the plasma sheet supports the *ring current* that flows in the magnetosphere during magnetic storms.

(d). At the edges of the magnetosphere, and obviously connected with the physics of the magnetopause, are *boundary layers*. Their composition and energy are governed by the solar wind and plasma in the magnetosheath.

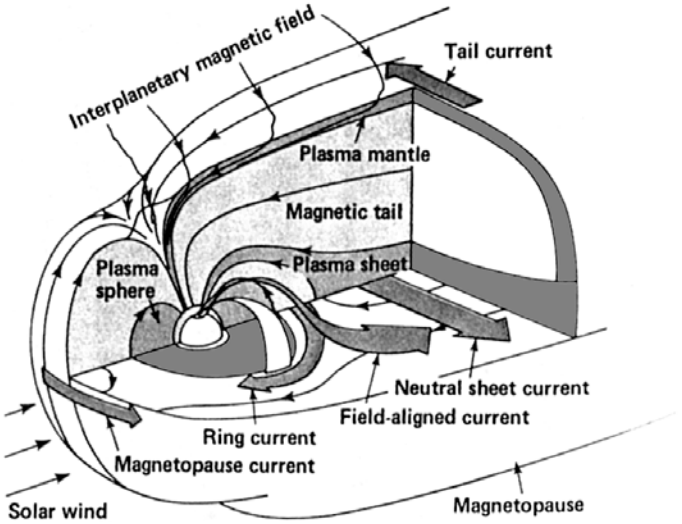


Figure 2.7. Plasma populations and current systems of the magnetosphere in three dimensions. (T. A. Potemra, *Johns Hopkins APL Tech. Digest* **4**, 276, 1983. © The Johns Hopkins University Applied Physics Laboratory, 1983. All rights reserved. Reproduced by permission.)

The locations of these particle populations are indicated in Figure 2.7. They are not merely incidental to the magnetosphere, but are in fact essential to its properties and behavior. Except for the boundary layers, each of these populations is discussed in a following section. In most of the magnetosphere the ratio of the energy density of the particles to that of the magnetic field (β) is less than unity, but there are exceptions.

Originally it was thought that most of the particles in the magnetosphere come from the solar wind, but, on the evidence of heavy ions observed in the magnetosphere, it is now recognized that there are also major sources in the ionosphere: specifically the auroral zones, the clefts, and the polar caps. It is thought that the solar wind is the dominant source in the distant magnetotail, but the ionospheric sources are important during storms and are sometimes dominant. (See also the discussion of the *polar wind*, Section 5.2.3.)

2.3.2 The plasmasphere

Ionized particles in the upper ionosphere (F region and topside) have temperatures up to several thousand degrees Kelvin, and electron energies are therefore several tenths of an electron volt. The particle density is typically $10^{10}\,\mathrm{m}^{-3}$ at 1000 km altitude, decreasing with increasing height – though not very rapidly because of the large scale height when atomic hydrogen is the principal atom. The theory of the protonosphere (Section 1.4.4) shows how ionospheric plasma flows up the field-lines to populate the protonosphere as far as the equatorial plane, provided

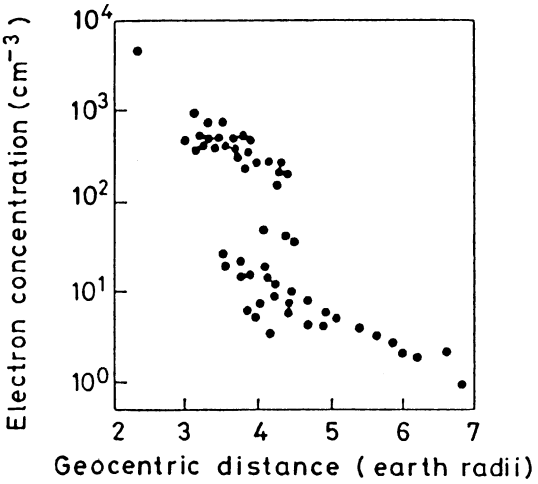


Figure 2.8. Electron density in the equatorial plane determined from whistlers. (J. A. Ratcliffe (after D. L. Carpenter), *An Introduction to the Ionosphere and Magnetosphere*. Cambridge University Press, 1972.)

that the field-lines are closed. Some of this plasma flows back to lower levels at night, where it helps to maintain the ionosphere during hours of darkness, but the plasmasphere nevertheless persists as a permanent feature of the inner magnetosphere. The outer boundary of the plasmasphere is called the *plasmapause*.

The plasmapause was discovered by a ground-based technique based on the reception of *VLF whistlers*. The whistler is a naturally occurring radio signal that propagates along the geomagnetic field between the northern and southern hemispheres. If the travel time of a whistler is displayed against frequency, it is seen that there is one frequency at which the travel time is a minimum. This is a characteristic of all whistlers. Not all show it clearly but those which do are called *nose whistlers*. The frequency corresponding to the minimum travel time indicates which field-line the whistler has traveled along, and the time taken can be interpreted to give the minimum electron density encountered along that field-line. A detailed discussion of whistlers and other magnetospheric noises would be beyond our scope; the reader who wishes to persue that interesting topic is referred to the book by Helliwell (1976).

By means of this technique it is possible to determine the variation of electron density in the equatorial plane, as in Figure 2.8. The remarkable feature of such plots is that they often exhibit a sudden drop in the electron density near $4R_E$. The decrease may be of a factor of ten or more within a distance of $0.5R_E$ or less. This edge is the plasmapause, sometimes also known as the *knee*. If it is traced inward along the geomagnetic field, it is found to correspond approximately to the ionospheric main trough which effectively marks the poleward extent of the mid-latitude ionosphere (Section 5.4). The plasmasphere thus occupies a doughnut-shaped region of the inner magnetosphere where the field-lines are not too

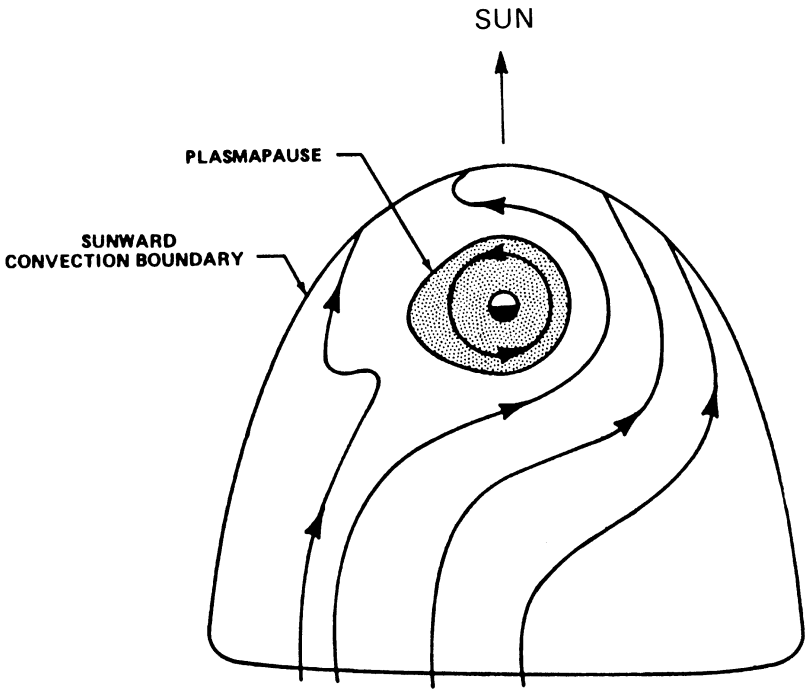


Figure 2.9. Plasma flow in the equatorial plane and the daily variation of the plasmapause. (After J. L. Burch, in *The Upper atmosphere and Magnetosphere*. National Academy of Sciences, Washington DC, 1977.)

distorted from the dipolar form, and the mid-latitude ionosphere is part of it.

The plasmapause is dynamic and variable. Its position varies during the day, the most notable feature being a bulge in the evening sector (Figure 2.9). In addition, the whole region contracts when geomagnetic activity increases, and there is then a gradual recovery over the next few days. Figure 2.10 shows measurements of the plasmapause position as a function of the global magnetic activity index K_p (described in Section 2.5.4). For most of the time it occurs between three and six R_E, though it has been detected as close to the Earth as $2R_E$ (i.e. only one R_E above the surface), and satellite data show occasions when no plasmapause was detected within seven or eight R_E.

According to whistler observations and *in situ* data (Carpenter and Anderson, 1992), the average geocentric distance to the plasmapause in Earth-radii (L_{pp}) is related to the greatest preceeding value of K_p ($\hat{K}_p$) by the empirical relation

$$L_{pp} = 5.6 - 0.46\hat{K}_p. \tag{2.8}$$

For this purpose $\hat{K}_p$ is derived over the previous 24 h, except that, for the hours 06–09, 09–12, and 12–15, respectively, the K_p values for the preceeding one, two, and three periods (of 3 h) are omitted. Carpenter and Anderson also give formulae for

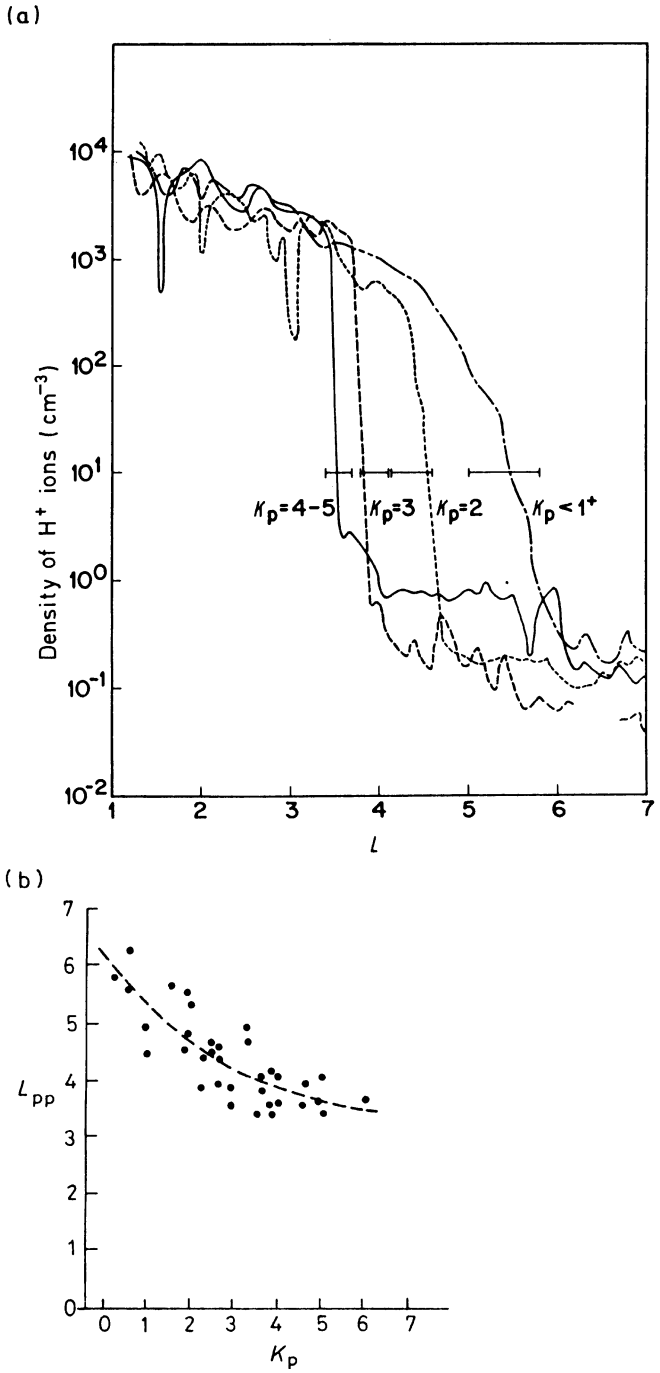


Figure 2.10. Variations of the plasmapause with the magnetic activity index, K_p. (a) Satellite observations of ion density, showing the plasmapause at several levels of Kp. (b) The relation between the plasmapause distance, L_{pp}, and K_p. (After C. R. Chappell *et al.*, *J. Geophys. Res.* **75**, 50, 1970, copyright by the American Geophysical Union.)

the electron density in the plasmasphere inside the pause and its regular variations, the thickness of the pause, and the distribution in the "trough" beyond it.

2.3.3 The plasma sheet

Beyond the plasmapause the electron density is much smaller and the temperature is much higher. Clearly this is a different population of particles. The electron and ion densities are each only about 0.5 cm^{-3}. Particle energies are generally 10^2–10^4 eV. The average energy of the electrons is about 0.6 keV and that of the protons about 5 keV. The total energy density of the particles is about 3 keV cm^{-3}.

The particular importance of the *plasma sheet* lies in its association with the central plane of the magnetotail where the magnetic field reverses. To a first approximation, the pressure of the particles in the sheet balances the magnetic pressure in the tail lobes. Thus,

$$nkT = B_T^2/(2\mu_0), \tag{2.9}$$

where B_T is the tail magnetic field outside the plasma sheet. As indicated in Figure 2.6, the plasma sheet follows the magnetic field down to lower altitudes in the vicinity of the auroral zone. It also continues round to the day side of the Earth. In the equatorial plane there is an identifiable, though variable, inner edge near $7R_E$ at midnight. The sheet is several Earth-radii thick (also variable) and it extends across the tail between the dusk and dawn sides. As the Earth's magnetic axis tilts seasonally and diurnally with respect to the Sun, the tail plasma sheet and neutral sheet oscillate north and south of the solar-ecliptic plane.

The magnetic field runs in opposite directions in the two lobes of the magnetotail, and the existence of a sheet of plasma between them creates an unusual physical situation. The configuration being far from dipolar, it represents a store of energy that could be released in the right circumstances. There is good evidence that a neutral line forms some $50R_E$ down the tail. Here the magnetic field is locally collapsed and plasma is accelerated both towards and away from the Earth. It is also known that events in the magnetotail are closely related to the phenomena of the aurora and the substorm, and it is thought that, at such times, a neutral line forms closer to the Earth, matter in the plasma sheet being then accelerated to higher energies. (This topic is discussed further in Section 6.4.)

2.3.4 Trapped particles

The discovery that there are energetic particles trapped in the magnetosphere was made early in the satellite era by Van Allen and colleagues at the University of Iowa. The information came from a Geiger counter, which they had built for the first successful satellite launched by the USA, Explorer 1, with the intention of studying cosmic rays. The cosmic rays were detected, but the high counting rates

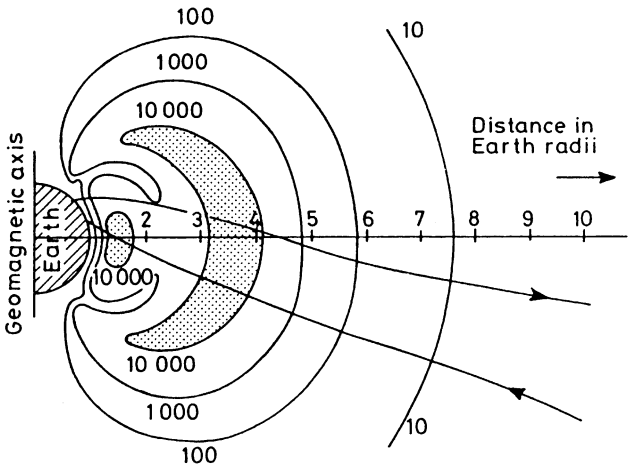


Figure 2.11. Van Allen's first map of the radiation belts, showing counting rates of the Geiger counter on Pioneer-3. (J. A. Van Allen and L. A. Frank, *J. Geophys. Res.* **64**, 1683, 1959, copyright by the American Geophysical Union.)

which were recorded in parts of the orbit indicated something much more exciting. Apart from its scientific value, this discovery was important in a political sense since it showed that the "space" near the Earth was not an empty void but contained at least some matter and energy and would, very likely, repay closer investigation. Figure 2.11 reproduces what is probably the most famous illustration of that period, showing the double structure deduced from the passage of Pioneer 3, an eccentric-orbit spacecraft that went out 107400 km in 1958. Thus were discovered the "inner" and "outer" Van Allen belts.

The division into two belts is something of an over-simplification because the structure of the trapping region depends on the nature and energy of the particles. The original discovery concerned protons with energy exceeding 30 MeV. Figure 2.12 shows a modern version of the trapped-particle distribution.

The mechanism of particle trapping is of interest. A trapped particle moves in three ways (Figure 2.13). It gyrates around a line of the geomagnetic field, bounces back and forth along the line of force between *mirror points*, and gradually drifts longitudinally around the Earth. The motions are based on a set of *adiabatic invariants*:

(1). the magnetic moment is constant;

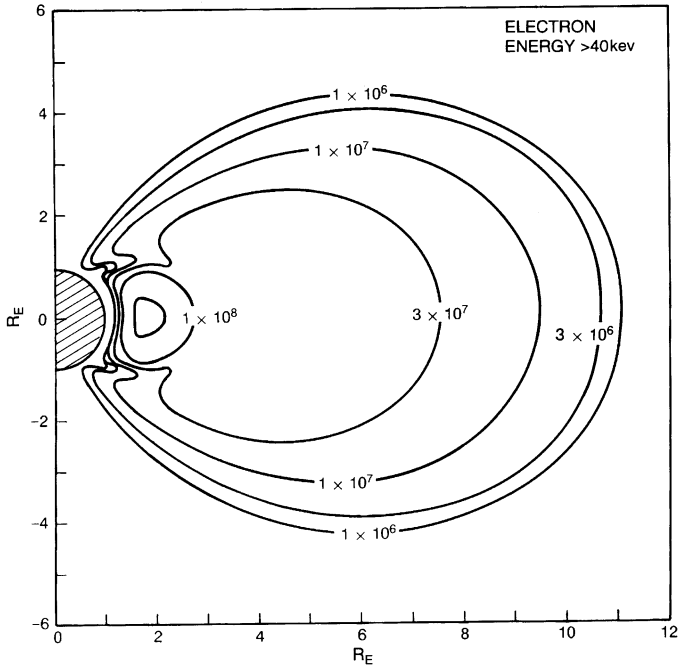
(2). the integral of the parallel momentum over one bouce between mirror points is constant; and

(3). the total geomagnetic flux enclosed by the drift orbit is constant.

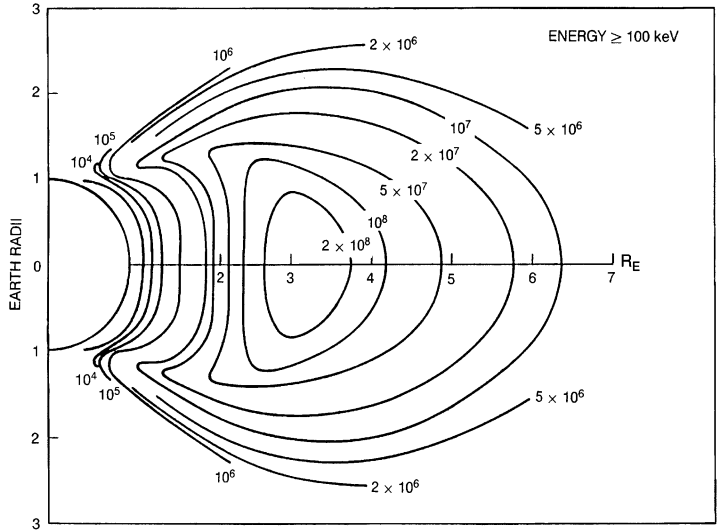
The first holds if the magnetic field does not change during a gyration period, the second if it does not change during a bounce period, and the third if it does not change during the time taken for the particle to encircle the Earth. Hence they are progressively less stringent.

The basic trapping mechanism is determined by the first invariant. Consider a charged particle gyrating in the geomagnetic field but also with a component of motion along the field. As the particle spirals from the equator towards the pole it

(a) Electrons



(b) Protons



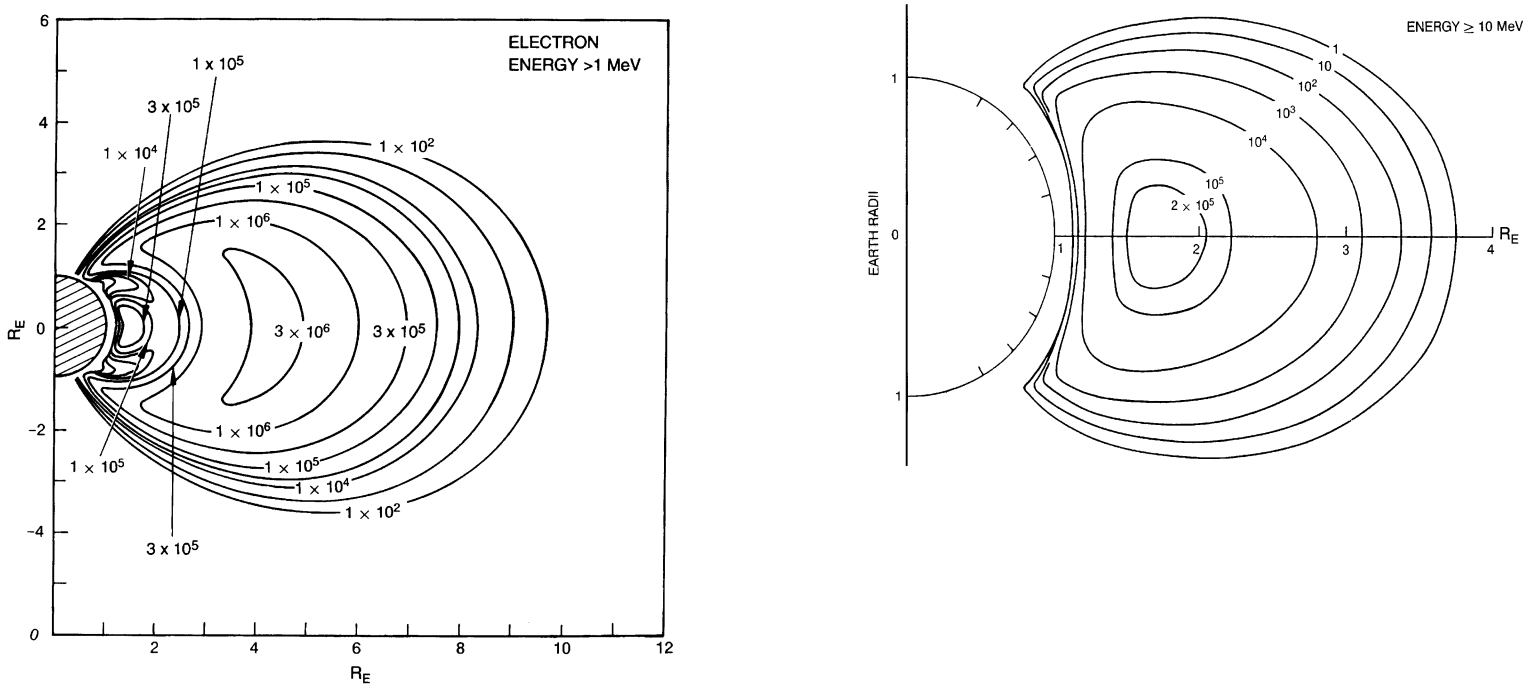
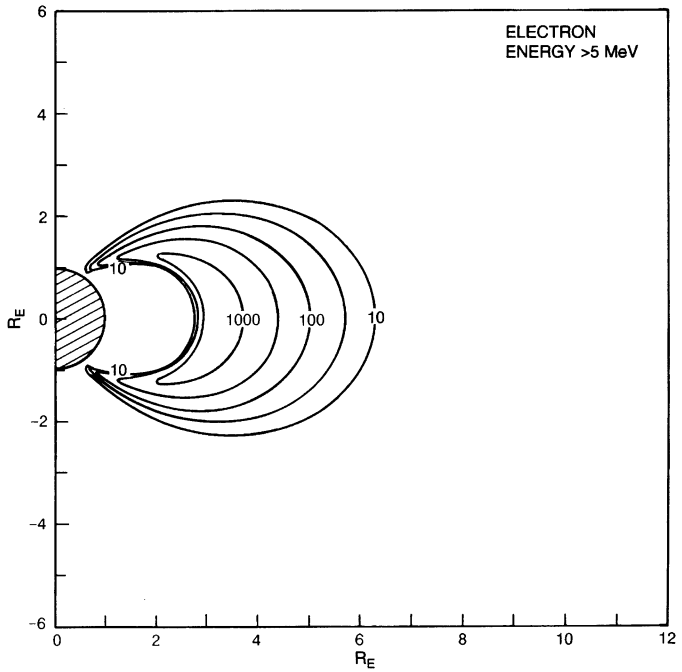


Figure 2.12. Spatial distributions of (a) trapped electrons of energy exceeding 40 keV, 1 MeV, and 5 MeV, and (b) trapped protons of energy exceeding 100 keV, 10 MeV, and 50 MeV. Since the particle flux falls with energy, these distributions are dominated by the particles just above the threshold stated. The fluxes are omnidirectional and in units of $cm^{-2}\ s^{-1}$. They are diurnal averages at sunspot maximum. The inner zone is seen in the fluxes of electrons of lower energy, and the *slot* between the zones is visible in the >1-MeV electron distribution. Otherwise the maximum occurs closer to the Earth at higher energy. (M. Walt, *Introduction to Geomagnetically Trapped Radiation*. Cambridge University Press, 1994.)

(a) Electrons



(b) Protons

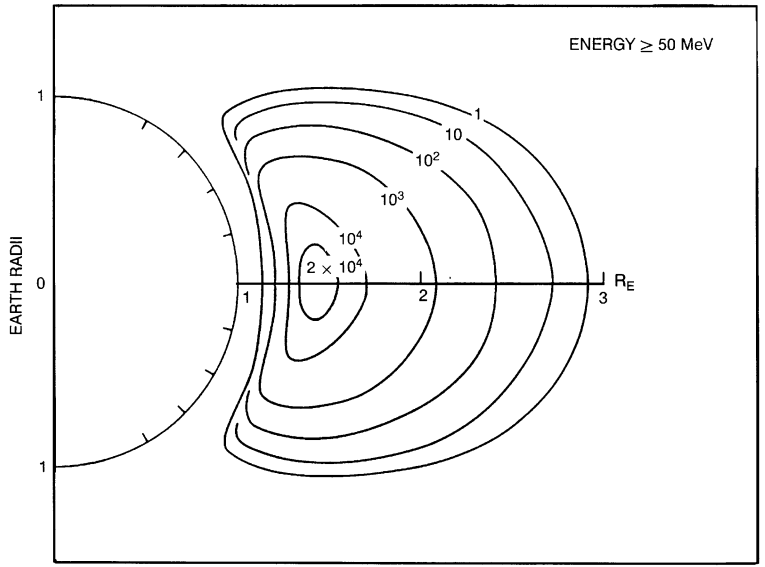


Figure 2.12. (*cont.*)

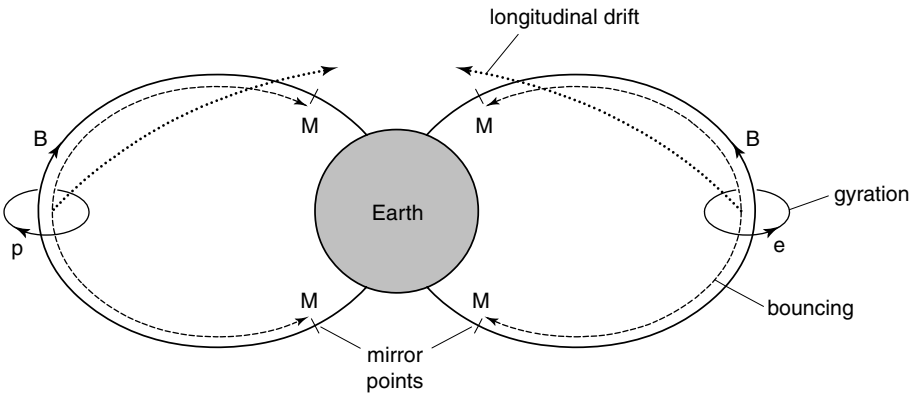


Figure 2.13. The motions of a particle trapped in the geomagnetic field.

moves into a region of stronger field. A consequence of the first invariant is that the component of kinetic energy perpendicular to the field is proportional to the magnetic flux density. Hence that component of the particle's energy increases, and the parallel component decreases by the same amount. Eventually, provided that the particle does not run into the atmosphere first, all the energy is transverse, forward motion stops, and, at that point, the *mirror point*, the particle is reflected back towards the equator and then into the other hemisphere. Since no energy is being lost or gained, the particle can continue thus for ever or until something else happens to it. The mirror point does not depend on the energy of the particle, but it is directly related to the particle's direction of travel (the *pitch angle*) as it crosses the equator.

If the mirror point is deep enough to be within the atmosphere, those particles will be lost. Correspondingly, there is, at any point along the path, some pitch angle within which all particles will be lost to the atmosphere at the next bounce. This defines the *loss cone*, which is generally a small angle of only a few degrees at the equator but increases progressively towards 90° as the Earth's surface is approached.

A trapped particle drifts longitudinally due to the forces which arise from the curvature and the radial change of intensity of the geomagnetic field. Electrons drift eastward and protons westward, and the rate of drift depends (more or less linearly) on the energy of the particle. The times which electrons of various energies and pitch angles take to orbit the Earth are illustrated in Figure 2.14.

The longitudinal drift path is determined by the second invariant. The principal population of trapped particles lives in the region of the magnetosphere where the field-lines are closed and almost dipolar. To remove particles from such orbits it is necessary to infringe one of the invariants. These particles are stably trapped. However, owing to the distorted form of the magnetosphere, some drift paths starting in the outer zone take particles into the magnetotail or into the solar wind. These particles cannot complete a full circuit of the Earth, and are trapped only temporally, in *pseudo-trapping* regions.

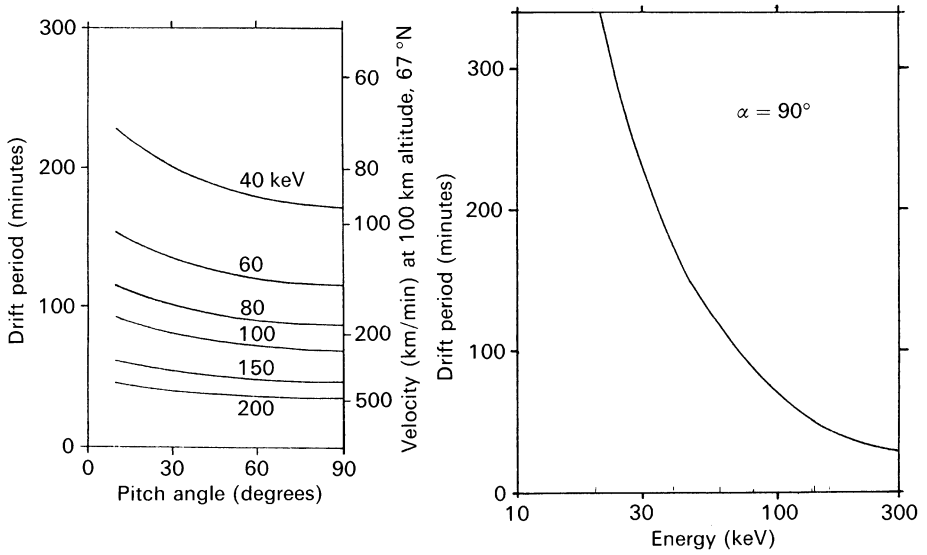


Figure 2.14. The time taken for trapped electrons to make one circuit of the Earth at the orbit of a geostationary satellite ($6.6R_E$): (a) as a function of pitch angle for various energies; (b) as a function of energy for pitch angle 90°. (a) Also gives the velocity of the footprint (at 67° latitude) 100 km above the Earth's surface. (P. N. Collis, private communication.)

The Van Allen particles do not have significant effects on the high-latitude ionosphere while they remain stably trapped. However, it is almost certain that the processes of trapping and loss, because they may transport energetic particles between different regions of the magnetosphere and then deposit them in the ionosphere, are important at high latitudes.

2.3.5 The ring current

One significant consequence of particle trapping is the formation of a ring current in the magnetosphere at or near the inner edge of the plasma sheet. The westward drift of trapped protons and the eastward drift of trapped electrons both contribute to a ring current directed clockwise as viewed from over the north pole. This current increases under disturbed conditions and may be detected with a magnetometer at the ground as the main phase of a magnetic storm (Section 2.5.2).

These particles are not the energetic ones typical of the Van Allen belts, but have been shown by direct measurement to be mainly protons of energy 10–100 keV. The current is generally located at a distance between four and six R_E (Figure 2.15) and its existence indicates that there is a concentration of charged particles in that region. Since the drift rate of a trapped particle is proportional to its energy, and all protons carry the same charge, it is easily shown that the total current in the ring is proportional to the total energy of the contributing protons.

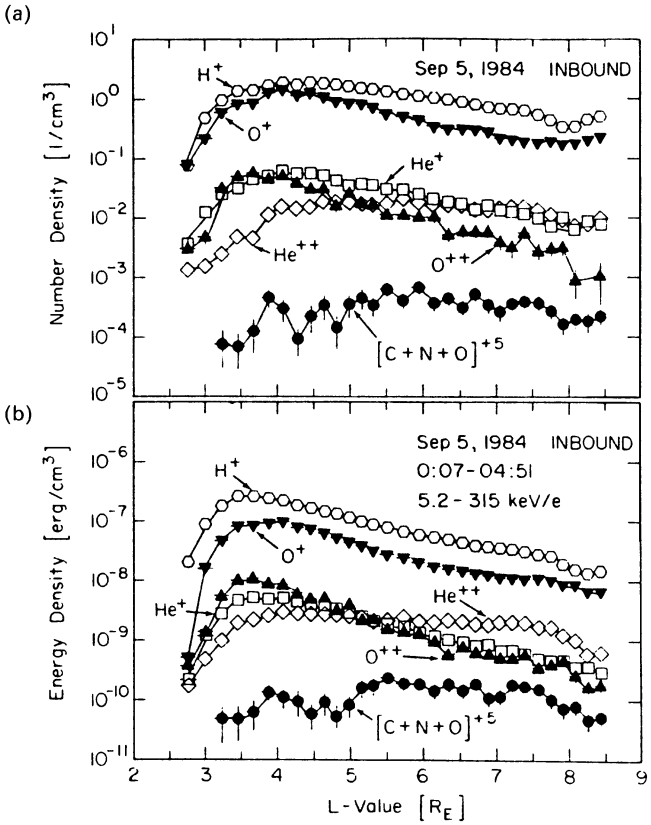


Figure 2.15. Radial profiles of various heavy ions during an inbound pass of the AMPTE satellite on 5 September 1984. (a) The number density. (b) The energy density of particles of energy 5–315 keV per unit electronic charge. (D. J. Williams, *Space Sci. Rev*. **42**, 375, 1985. With kind permission from Kluwer Academic Publishers; after G. Gloeckler *et al.*, private communication.)

It is possible to have a ring current, or a component of a ring current, which goes only part way around the Earth: a *partial ring current*.

2.3.6 Birkeland currents

Since current has to be continuous, we may well ask how the circuit of a partial ring current is completed. Since the early 1970s it has become generally accepted that currents may flow along the magnetic field-lines between the magnetosphere and the ionosphere. Such currents had first been suggested in 1908 by K. Birkeland, but direct measurements *in situ* were required in order to prove their existence. Typical distributions of *Birkeland currents*, sometimes called *field-aligned currents*, are illustrated in Figure 2.16.

The currents fall into several distinct regions, and by convention the poleward one is "region 1" and the equatorward one is "region 2", irrespective of the

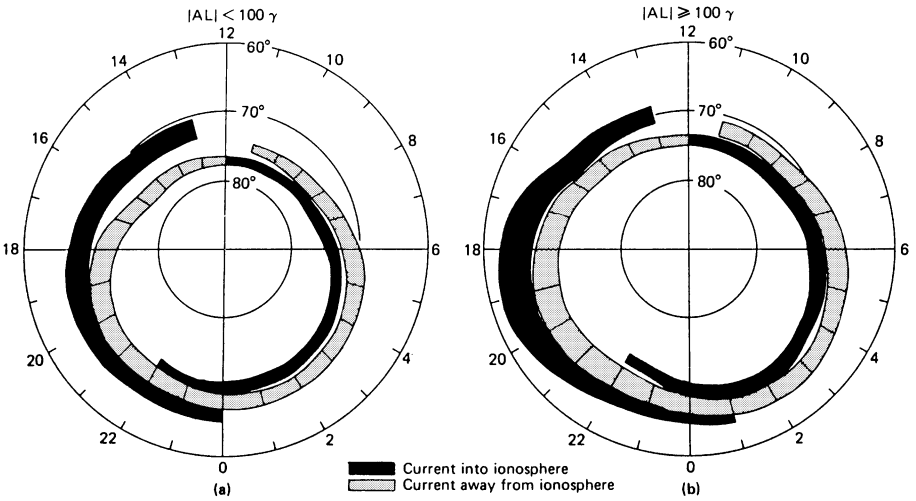


Figure 2.16. Distributions of Birkeland currents during (a) weak and (b) active disturbances. (T. Iijima and T. A. Potemra, *J. Geophys. Res*. **83**, 599, 1978, copyright by the American Geophysical Union.)

direction of the current. The intensity varies during the day in each region and is generally up to 1 or 2 μA m^{-2}. The total field-aligned current is 10^6–10^7 A.

The concept of the Birkeland current has profoundly affected ideas about current systems in the ionosphere and magnetosphere. Magnetic observations made at the surface of the Earth may always be interpreted as a two-dimensional current system flowing horizontally at some (unspecified) height in the ionosphere, and the earlier analyses were always performed in this way. These are actually *equivalent current systems*, which are mathematically correct but are not the only solutions if vertical current is also allowed. The inclusion of Birkeland current has led to distributions that include both ionospheric and magnetospheric parts, and are physically more instructive.

2.4 The dynamics of the magnetosphere

2.4.1 Circulation patterns

A static description of the magnetosphere is alright as far as it goes, but there are certain facts and phenomena that it cannot hope to explain. If we perform a simple calculation of the shape of the magnetopause from the pressure of the solar wind (as suggested in Section 2.2.3), we are using just the component of force normal to the boundary. However, if the solar wind also exerts a force tangential to the boundary, energy will be transfered into the magnetosphere from the solar wind, and we have the possibility that material within the magnetosphere will be set in motion. The situation is then somewhat akin to that of a falling raindrop,

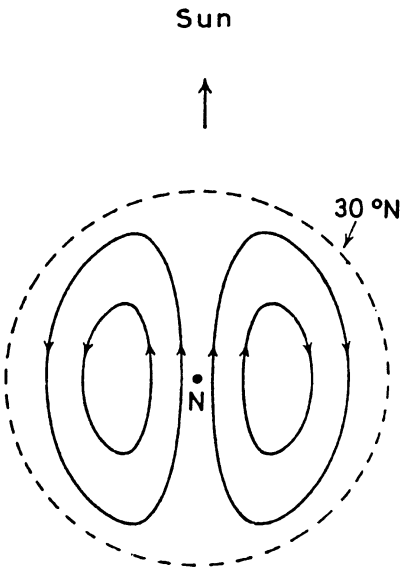


Figure 2.17. The S_q^p current system in the polar regions due to the circulation of magnetospheric field-lines. (J. A. Ratcliffe, *An Introduction to the Ionosphere and Magnetosphere*, Cambridge University Press, 1972.)

in which liquid is swept back at the surface and returns down the middle of the drop.

The concept of a circulating magnetosphere driven by viscous interaction at the surface was put forward by Axford and Hines in 1961. The nature of the viscous interaction was not specified but was thought to be effectively some kind of friction. Experimental support for circulation came from a study of the S_q^p *current system*, whose existence may be inferred from observations with magnetometers at medium and high latitudes. The current system causing S_q^p – the term means the polar part (p) of the magnetic variation related to the solar day (S) which is observed under magnetically quiet conditions (q) – is illustrated in Figure 2.17. The current flows over the poles from night to day, and there are return currents at lower latitudes. We show the pattern here because it is one that will prove basic in later considerations of the high-latitude ionosphere.

Since electrons are tied to the magnetic field in the dynamo region (whereas the positive ions move with the neutral air), the S_q^p current flow can be interpreted as a motion of magnetic field-lines in the opposite direction; that is, over the pole from the day to the night side of the Earth. In the magnetosphere, therefore, the field-lines circulate over the poles from the day to the night sectors of the Earth with a return flow around the dawn and dusk sides.

Figure 2.18(a) shows the basic pattern of magnetospheric circulation in a section through the equatorial plane. Figure 2.18(b) includes the distortion due to the Earth's rotation which carries the inner part of the magnetosphere with it. Section 2.4.4 indicates how the combined effect of two circulation patterns may be handled.

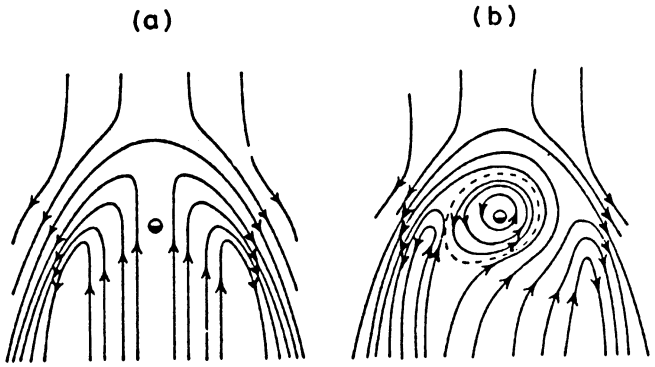


Figure 2.18. Patterns of magnetospheric circulation in the equatorial plane: (a) if due to "friction" of the solar wind at the magnetopause; and (b) if the Earth's rotation is included. (After A. Nishida, *J. Geophys. Res.* **71**, 5669, 1966, copyright by the American Geophysical Union.)

Present evidence, however, is that, although viscous interaction plays some part in driving the magnetosphere, it is not the major cause. One reason is that the solar wind is so tenuous (having a mean free path between collisions of perhaps 10^9 km!) that it is hard to believe in a sufficient amount of friction at the magnetopause. Attention therefore moved to an alternative mechanism, based on the work of J. W. Dungey concerning interconnection between the interplanetary magnetic field and the geomagnetic field. The various field configurations that arise when a dipole is placed in a uniform magnetic field are easily illustrated in simple laboratory experiments using a bar magnet situated in an external field. When the fields are parallel there are neutral points on the equator and connections between the two fields. When they are antiparallel, the neutral points are over the poles and there is no interconnection. Figure 2.19 depicts a distorted dipole field representing the geomagnetic field in polar section, with the addition of (a) a northward IMF and (b) a southward IMF. In the second case, neutral points are formed in the equatorial plane and some lines of the IMF connect to geomagnetic lines. This is not so in the first case. We have seen that the IMF tends to lie in the solar-ecliptic plane, oriented at the "garden-hose" angle, but there is usually a north–south component as well and it is this, when it is directed southward, which connects with the geomagnetic field.

The IMF is frozen into the solar wind and is therefore carried along with it. When geomagnetic field-lines are connected to those from the IMF they are dragged over the poles from the sunward neutral point, as in Figure 2.19(c), and thereby transported from the day to the night side. While they are over the polar caps the field-lines are open in the sense that they do not connect back to the other hemisphere in any simple or obvious manner. In the tail these lines reconnect and move back towards the Earth. The above picture is of course a simplified one. Detailed consideration taking account of the three-dimensional form of the magnetosphere shows how it is possible to have a degree of connection when the IMF

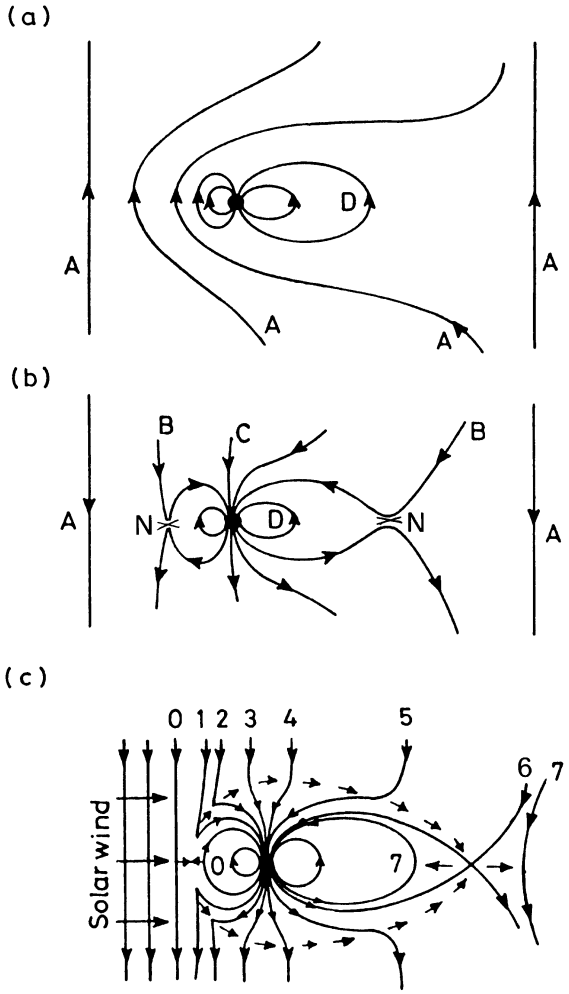


Figure 2.19. The interaction of terrestrial and interplanetary magnetic fields seen in polar section: (a) northward IMF; (b) southward IMF; and (c) circulation due to the flow of the solar wind. (After C. T. Russell, *Critical Problems of Magnetospheric Physics*, 1972, after J. W. Dungey, 1963, and R. H. Levy *et al.*, *Am. Inst. Aeronaut. Astronaut. J.* **2**, 2065, 1964.) A: Interplanetary field-line. B: Interplanetary field-line connecting to, or disconnecting from, a geomagnetic field-line. C: Open geomagnetic field-line. D: Closed geomagnetic field-line. N: Neutral point. 0–7: Successive positions of a selected interplanetary field-line.

is northward, and how the east–west component affects the connection point and the resulting circulation pattern. Also, it is thought that viscous interaction does make some contribution; a minor one when the IMF has a southward component, but perhaps the main one when the IMF is northward (and the circulation is then much reduced).

The details and mechanisms of magnetic connection and magnetospheric circulation continue to be topics for research, but there is little doubt that the IMF plays an important role. Signatures indicating magnetic connection are observed

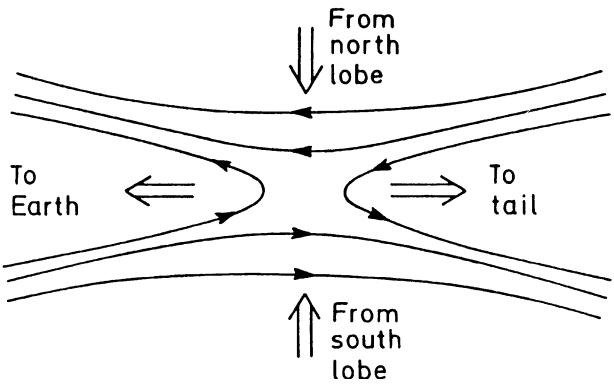


Figure 2.20. An X-type neutral line in the magnetotail. Plasma flows in from the north and south lobes, and leaves Earthward and tailward along the plasma sheet.

by spacecraft passing through the magnetopause, and the level of geophysical activity increases when the IMF has a southward component. At such times (a) substorms (Section 6.4) occur more frequently; (b) the magnetic flux in the tail lobes is increased; (c) the auroral zone and the dayside cusps are displaced equatorward; and (d) the dayside magnetopause moves inward – all suggesting that a southward IMF increases the rate of magnetospheric circulation.

Magnetospheric circulation is a concept of great significance, not only in magnetospheric theory but also, as we shall see, for the high-latitude ionosphere.

2.4.2 Field merging

Magnetospheric circulation requires that, on both the day and the night sides of the Earth, magnetic field-lines are broken and then reconnected in a different configuration. The simplest model of such a process, the *X-type neutral line*, is illustrated by Figure 2.20. This shows a situation in the central region of the tail. The configuration cannot be static because the tension in the field-lines will produce net forces towards the Earth and into the tail. However, there can be dynamic equilibrium, in which the depletion is replaced by other field-lines moving in from the lobes. Those lines are, of course, replaced by others moving over the poles from the day side of the Earth. There can also be a *Y-type neutral line*, where the field continues to converge on the tailward side; in that case all the reconnected lines move towards the Earth.

The theories of magnetic reconnection came originally from studies of solar flares. In the magnetosphere the process is thought to be that of *fast reconnection*, first proposed by Petschek (1964) more than 30 years ago. This mechanism invokes an Alfvén wave, which allows reconnection to proceed more rapidly than would diffusion only. The velocity of reconnected field-lines towards the Earth is estimated as about 100 km s^{-1}, and the drift towards the neutral sheet as about 10

km s^{-1}. Particles on the field-lines passing through the reconnection region are accelerated in the direction of the contraction. It is likely that reconnection in the tail occurs not steadily but intermittently in limited regions, and this is probably important in the causes of the aurora.

While a pattern of circulation must plainly include reconnection on the night side, it is the connection between the IMF and geomagnetic field-lines on the day side which drives the circulation. Though various ideas have been suggested, the details of this process have not been finally decided. Obviously, a geomagnetic flux tube has to break and connect with an IMF tube, and this is the event which has been identified from the magnetic signature recorded by a nearby spacecraft, as a *flux-transfer event (FTE)*. The newly connected tube of plasma then moves poleward into the boundary layer and joins the general circulation. FTEs are frequent, though individually of short duration and limited spatial extent ($(0.5–1)R_E$). There are more FTEs when the IMF has a strong southward component, and none when it is northward. Presumably, details also vary with the direction of the east–west component. "Quasi-steady" connection is also a possibility.

2.4.3 Magnetospheric electric fields

It is sometimes helpful to regard the dynamic magnetosphere in terms of an electric dynamo and a motor. The magnetosphere may be treated as a magneto-hydrodynamic generator, in which a jet of plasma (the magnetosphere) is forced through a static magnetic field (the IMF) and an electric potential is developed by dynamo action. The total potential difference across the magnetosphere is

$$V_T = vLB_n, \tag{2.10}$$

where v is the solar-wind speed, L is the width of the magnetosphere, and B_n is the magnetic flux density normal to the boundary. Its value is estimated as about 60 kV, equivalent to an electric field of about 0.3 mV m^{-1}. The electric field is directed from the dawn to the dusk side of the magnetosphere. The same potential difference appears across the open field region of the high-latitude ionosphere, the field again being directed from dawn to dusk.

The motion of magnetoplasma within the magnetosphere can now be regarded as the effect of this electric field on the geomagnetic field as in an electric motor, according to

$$\mathbf{v} = \mathbf{E} \times \mathbf{B}/|\mathbf{B}|^2, \tag{2.11}$$

where $\mathbf{v}$ is the velocity, $\mathbf{E}$ the electric field, and $\mathbf{B}$ the magnetic field. The magnitude is simply

$$v = E/B. \tag{2.12}$$

The potential distribution across the magnetotail maps along the field-lines into the polar caps, where it is more accessible to direct measurement, and relationships have been found with the speed of the solar wind and the magnitude and direction of the IMF (Section 5.1.2). If the potential difference across the polar cap is 60 kV, the field-line velocity is about 300 m s^{-1}.

2.4.4 The dynamics of the plasmasphere

A good example of treating the dynamics of the magnetosphere in terms of electric fields is the question of the location of the plasmapause, the boundary between the plasmasphere and the outer magnetosphere. The higher levels of the plasmasphere are created by ionospheric plasma moving up and down closed geomagnetic field-lines. However, to explain the dynamics of the plasmasphere as a whole, it is also necessary to take account of the circulation of the magnetosphere. The inner magnetosphere co-rotates with the Earth while the outer magnetosphere follows its own circulation pattern under the control of the solar wind. Generally speaking, the plasmasphere exists on the co-rotating field, and the plasmapause marks the boundary between the inner and outer regions.

If we imagine that the plasmasphere is observed by a person fixed in space (i.e. not rotating with the Earth), we can show that its motion in the equatorial plane may be ascribed to a *co-rotation electric field* of magnitude

$$E_r = BLR_E\omega, \tag{2.13}$$

where B is the geomagnetic flux density, L is the observer's geocentric distance in Earth-radii, R_E is the radius of the Earth and ω is the Earth's angular velocity. The plasmapause occurs approximately where the cross-tail and co-rotation fields are equal:

$$E_T = \frac{B_E}{L^3} LR_E\omega, \tag{2.14}$$

where B_E is the geomagnetic flux density at the surface of the Earth at the equator, and we have used the radial variation of flux density in a dipole field, $B = B_E/L^3$. The condition expressed by Equation (2.14) marks the transition between the circulation regimes of the inner and outer magnetosphere. Putting in numerical values gives:

$$L^2 = 14.4/E_T \text{ (mV m}^{-1}\text{)}. \tag{2.15}$$

If the tail field is 1 mV m^{-1} we expect to find the plasmapause at about $4R_E$.

A computation of plasma convection about the Earth was shown in Figure 2.9. In general these flow lines are also equipotentials. The bulge in the plasmasphere in the evening sector, a well-established feature, occurs because the co-rotation and cross-tail fields are in opposite directions on the evening side.

We now see that the principal dynamics of the plasmasphere are (1) filling and emptying along the tubes of force from the ionosphere, which depends on the time of day; and (2) rotation about the Earth in a pattern that is also affected towards its outer edge by the circulation of the magnetosphere. The second factor explains why the location of the plasmapause varies with geomagnetic activity. An increase in the circulation of the magnetosphere implies that condition (2.15) is satisfied closer to the Earth and the plasmasphere must then be smaller. It is thought that the change of circulation peels off layers of plasma, which may exist as detached regions before becoming lost to the outer magnetosphere or into the solar wind.

When activity returns to normal the magnetospheric circulation and electric fields return to their previous state, but now the outer tubes of magnetic flux are devoid of plasma. These gradually refill from the ionosphere over a period of several days. The rate of filling is determined by the rate of diffusion of protons from the upper ionosphere (where they are formed by charge exchange between hydrogen atoms and oxygen ions – Section 1.4.4), and by the volume of the flux tube to be filled. Since the latter varies as L^4 it takes much longer to refill tubes originating at higher latitude, and, since active periods may recur every few days, there will be periods when the outer tubes are never full. It is probably safe to say that the plasmasphere always suffers from some degree of depletion.

2.5 Magnetic storms

2.5.1 Introduction

The ionospheric storm was introduced in Chapter 1, but the *magnetic storm*, the main part of which is due to the ring current, is probably the more fundamental. Like the ionospheric storm, it may last from a few hours to several days and it often exhibits three phases. It has been known – though not by its present name – since the eighteenth century from its effect on a compass needle, but progress in understanding any of the storm phenomena dates only from modern times. Because magnetic storms can be monitored without great difficulty using a magnetometer, and long runs of such measurements exist, the magnetic storm has come to be a common reference point in geophysical studies.

Although there are superficial similarities between magnetic and F-region storms, the physical connections are not so obvious. These are phenomena that have not proved amenable to simple explanations, and some major questions remain. Part of the problem is that a chain of events is involved. The primary cause is almost certainly the solar wind, affecting the magnetosphere. The magnetospheric consequences then affect the upper atmosphere, and, in some cases, there might even be contributions ascending from the troposphere or stratosphere.

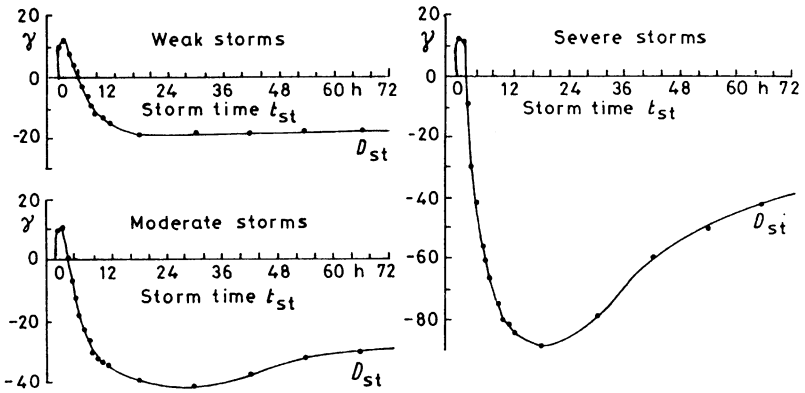


Figure 2.21. Typical magnetic storms registered by an equatorial magnetometer. (After M. Sugiura and S. Chapman, *Abhandl. Akad. Wiss. Gottingen Math.-Phys. Kl.*, Special Issue 4, 1960.)

2.5.2 The classical magnetic storm and the D_{st} index

The typical magnetic storm is illustrated in Figure 2.21. Like the ionospheric storm, this *classical magnetic storm* consists of three phases:

(a). an increase of magnetic field lasting a few hours only, called the *initial phase*;

(b). a large decrease in the H component building up to a maximum in about a day, the *main phase*; and

(c). a slow recovery to normal over the next few days, the *recovery phase*.

The initial phase is caused by the compression of the front of the magnetosphere with the arrival of a burst of solar plasma, as in Chapman and Ferraro's theory of 1930 (Section 2.2.2). The main phase is due to the ring current which was introduced in Section 2.3.5. The recovery phase is simply a recovery to the pre-storm condition as the ring current decays.

The D_{st} *index* of magnetic storms is derived from low-latitude magnetograms. In units of nanoteslas ($=\gamma$), it simply expresses the reduction of the magnetic H component at the equator due to the ring current, and it serves as a useful indicator of the intensity and duration of individual storms. If we assume a distance for the ring current, its magnitude may be derived from the equation

$$\Delta B \approx 3\pi I/(10r) \approx Ir, \tag{2.16}$$

where ΔB is the change of magnetic flux (in nanoteslas), I is the current (in amperes) and r is the assumed distance (in kilometers). Equation (2.16) is derived from the standard formula for the flux density at the center of a current loop, but with a correction for currents induced in the ground.

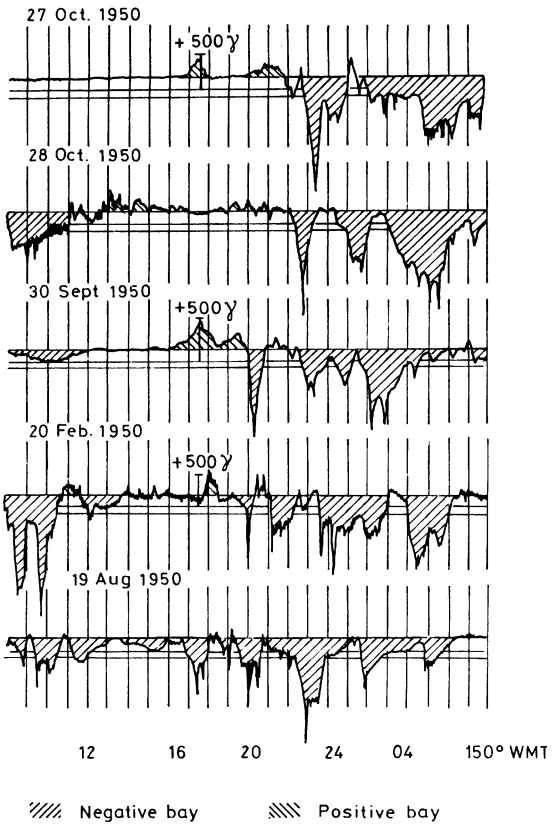


Figure 2.22. Examples of positive and negative magnetic bays recorded at College, Alaska. The time zone is that of the 150° W meridian. (After S.-I. Akasofu, *Polar and Magnetospheric Substorms*, Reidel, 1968, with kind permission from Kluwer Academic Publishers.)

2.5.3 Magnetic bays at high latitude; the auroral electrojet

The magnetic storm appears in a different guise at high latitude. By contrast with records from low latitudes, where the effects are due to the growth and decay of the ring current in times measured in hours and days, magnetograms recorded in and near the auroral zone show more rapid changes. The typical pattern there is a series of *bay* events with typical durations of tens of minutes to an hour or two, such as those illustrated in Figure 2.22. The magnitude of the perturbation in the horizontal component (H) can be as much as 1000 nT, and its sign tends to be positive before midnight and negative afterwards. Where the sign changes is called the *Harang discontinuity*.

The magnetic bay is caused by an electric current, the *auroral electrojet*, flowing not in the magnetosphere but in the auroral E region. To explain the sign of the bay, the current flow must be eastward before midnight and westward afterwards: i.e. converging on the midnight meridian. Obviously there must also be return currents for continuity. Chapman's original interpretation assumed that the currents only flowed only horizontally and this gives a pattern called the *SD current system*, which is composed of two electrojets with return currents at higher and lower latitudes as in Figure 2.23. This pattern was obtained by averaging the daily

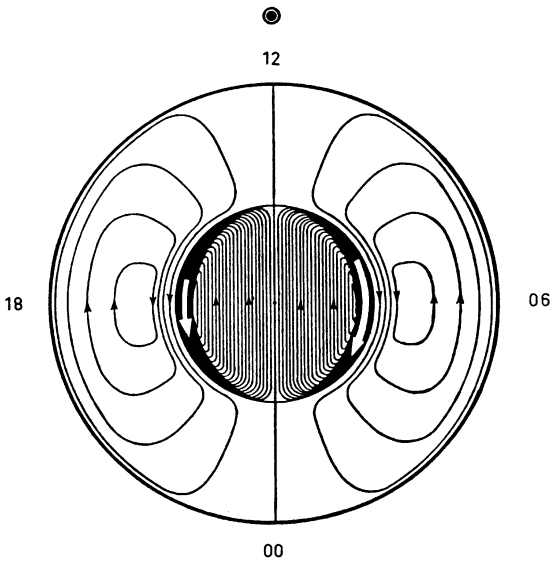


Figure 2.23. Chapman's original SD current system. The SD analysis takes magnetic disturbance vectors observed simultaneously at a number of stations and infers a current system that could give rise to them. (S.-I. Akasofu, *Polar and Magnetospheric Substorms*, Reidel, 1968, with kind permission from Kluwer Academic Publishers.)

magnetic variations during the first two days of a number of storms, a proceedure that concentrates the inferred current into the auroral zone. If a three-dimensional current system is allowed, other distributions become possible; modern interpretations include Birkeland currents (Sections 2.3.6 and 6.4.4).

2.5.4 Magnetic indices

The magnetic bays of the kind illustrated in Figure 2.22 are the basis of several magnetic indices, which are regularly compiled and published. The primary purpose of these indices is to quantify the intensity of geomagnetic disturbance and thereby provide a common reference point and a basis for comparison between different observations. The bays, and of course the electrojets causing them, are part of the substorm phenomenon – see Section 6.4 – and as such may be expected also to be related to the intensity of substorms and their frequency of occurrence. The most useful indices are probably those known as K_p, A_p, and AE.

K_p and A_p

K_p is based on the range of variation within 3-h periods of the UT day observed in the records from about a dozen selected magnetic observatories. After local weighting, and averaging, the K_p value for each 3 h of the day is obtained on a scale from 0 (for "very quiet") to 9 (for "very disturbed"). The scale is quasi-logarithmic, and the integer values are sub-divided into thirds by use of the symbols + and −: thus, 2, 2+, 3−, 3, 3+, etc.

A_p is a daily index, obtained from the same basic data, but converted to a linear scale (the 3-h a_p) and then averaged over the U. T. day. The value of the intermediate a_p is approximately half the range of variation of the most disturbed mag-

netic component measured in nanoteslas. The relationship between K_p and a_p is given in Table 2.1.

It is convenient to present K_p as a *Bartels musical diagram*, as in Figure 2.24, in which form it often shows how the activity tends to recur with the 27-day solar rotation. In such cases the diagram has some predictive value, but the 27-day recurrence is not always in evidence, being more apparent during the declining phase of the solar cycle.

Early measurements of the solar wind led to an empirical relation between the speed of the solar wind and K_p:

$$v\ (\mathrm{km\ s^{-1}}) = (8.44 \pm 0.74)\Sigma K_p + (330 \pm 17), \tag{2.17}$$

where ΣK_p is the sum of the eight K_p values over a U. T. day. This was an important early result in that it demonstrated a relationship between the solar wind and disturbances of the geomagnetic field.

AU, AL, and AE

The magnetic observatories which contribute to K_p and A_p are situated at various latitudes and longitudes, but with a preponderance in the higher middle latitudes, i.e. the equatorward side of the auroral zone. To achieve an index more tightly related to the auroral regions and to provide better time resolution, AU, AL, and AE were invented by Davis and Sugiura (1966). These indices are obtained by a rather different proceedure. Magnetograms from observatories at several different longitudes around the auroral zone are superimposed and the upper and lower envelopes are read. The upper envelope is AU, the lower envelope is AL and the difference between the envelopes is AE. AL indicates the greatest positive excursion in the auroral zone (probably a pre-midnight bay), AU the greatest negative

Table 2.1. *The relationship between* K_p *and* a_p

K_p	a_p
0	0
1	3
2	7
3	15
4	27
5	48
6	80
7	140
8	240
9	400

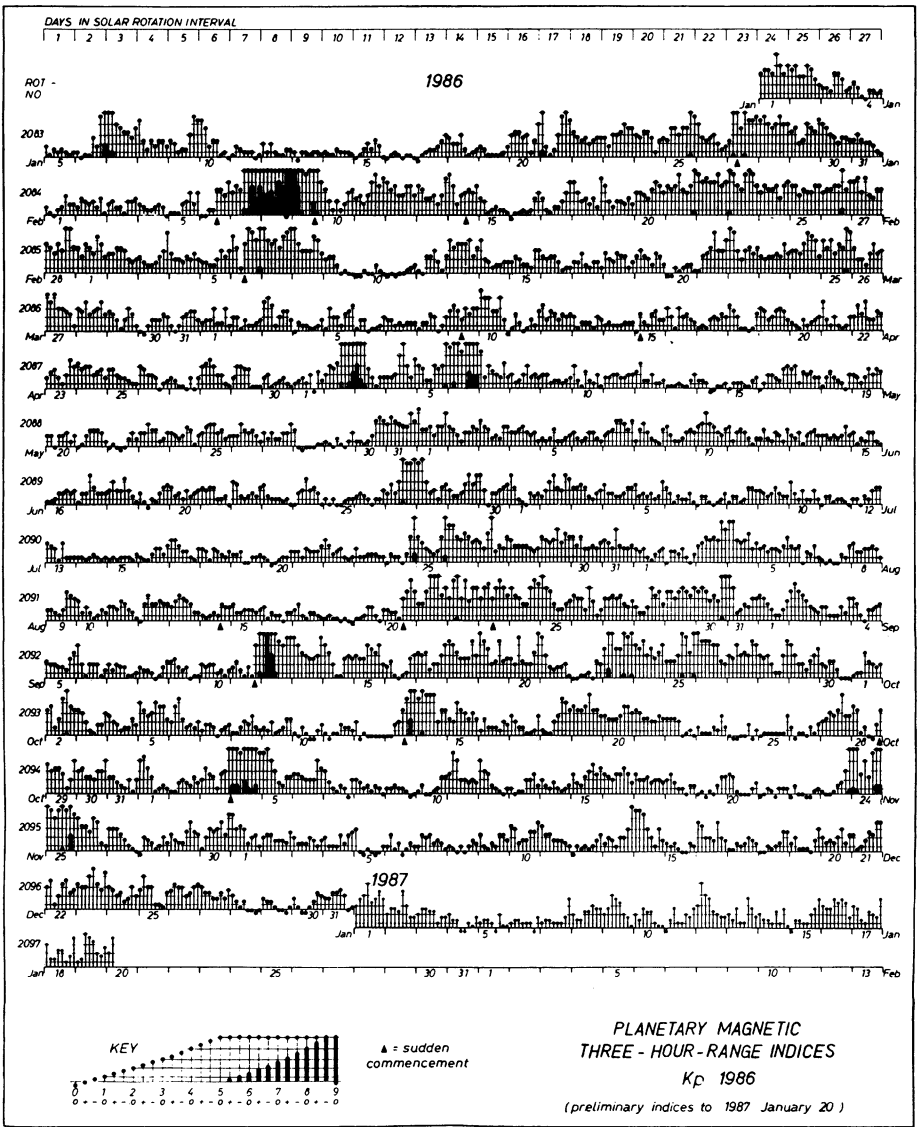


Figure 2.24. A Bartels diagram of K_p. (*Solar–Geophysical Data*. National Geophysical Data Center, NOAA, Boulder, Colorado, February 1987.)

excursion (probably post-midnight), and AE, which is the most widely used of the three, is taken as a general indicator of auroral-zone activity irrespective of the local time. Figure 2.25 shows an example. The mean of AU and AL is plotted as A0. The values are published at an hourly interval in printed reports, and they are available at a 2-min interval by special request.

In principle, AU is a measure of the eastward auroral electrojet and AL a measure of the westward electrojet. However, both of these indices may be affected by the ring current. The advantage of AE, being their difference, is that it depends

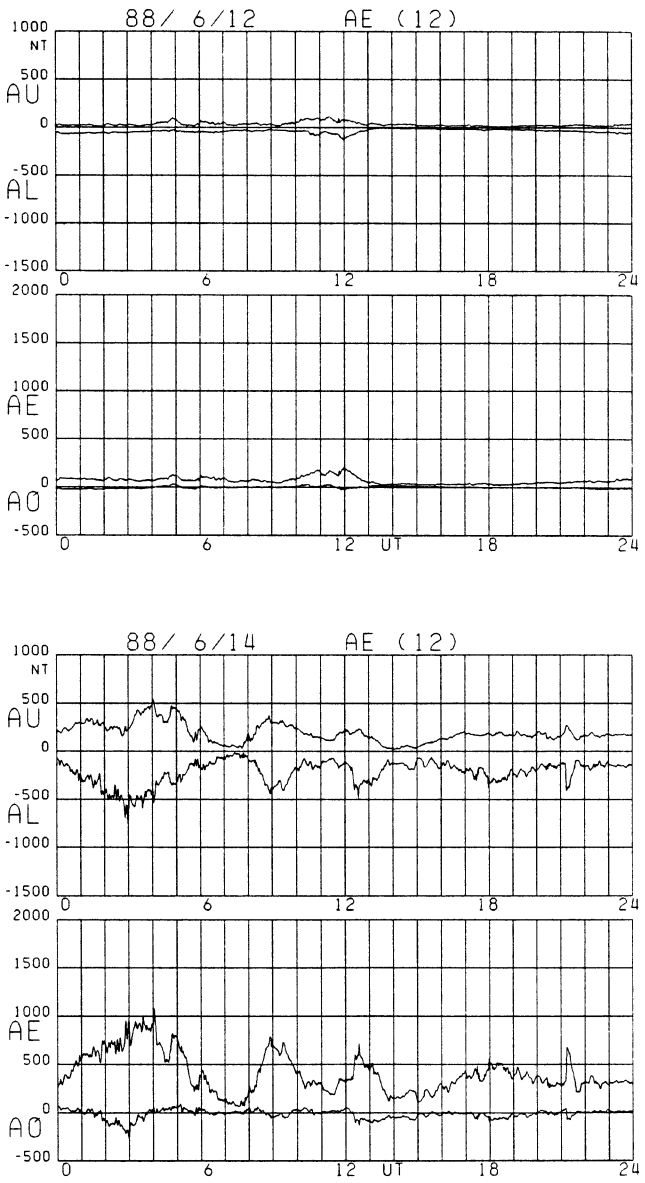


Figure 2.25. AU, AL, AE, and AO indices for 12 and 14 June 1988, the first a quiet day ($A_p = 4$) and the second a disturbed one ($A_p = 20$). (*Data Book 23*, September 1994, World Data Center-C2, Kyoto University, Kyoto.)

solely on the eastward and westward electrojets and should be independent of the ring current and any other zonal current. AE is particularly valuable for indicating when a magnetic substorm occurs. It is also well correlated to the energy coupled into the magnetosphere from the solar wind (Section 6.4.6).

Indices such as AE are obviously more sophisticated than K_p, and their preparation requires a correspondingly greater effort, so that the values may not become generally available for a year or more. (K_p, on the other hand, can be obtained through the World Data Centers within a few days.)

The history of magnetic indices and their derivations, advantages, and disadvantages are discussed in detail by Maynaud (1980). Some of the older indices continue to be of interest and are still produced to maintain continuity. The C index, one of the first, is a simple character index in which 0, 1, and 2 mean simply "quiet", "moderately disturbed" and "disturbed", respectively. The R and Q indices are range indices like K, but are derived at hourly and 15-min intervals, respectively, instead of every 3 h.

We shall refer frequently to magnetic indices of one variety or another when dealing in later chapters with the behavior of the disturbed ionosphere at high latitude. They are the common currency of geophysical disturbance, and are useful beyond the limited topic of magnetic disturbance because of their ease of measurement and the long runs of values accumulated over the years.

2.5.5 Great magnetic storms and a case history

Some storms are so intense and their effects so dramatic that they attract both scientific and popular attention. Though such storms are rare, they serve to illustrate how great the effects can be in extreme cases. Table 2.2 gives the top ten magnetic storms of modern times, ranked in order of the maximum A_p occurring during the storm. In terms of the equatorial index D_{st}, which measures the strength of the ring current (and comes out with a negative value, since the H component is reduced), the greatest storm was that of 13 March 1989. (Values are not available for the earliest storms of Table 2.2 because D_{st} has been derived only since 1957.)

From more extended lists of great storms it has been noted that (1) most of them occurred after solar maximum rather than at the maximum or before it; and (2) more than half occurred during the four months of the year nearest the equinoxes: that is, during March, April, September, and October.

Table 2.2. *The top ten magnetic storms of modern times (after J. A. Joslyn, private communication)*

Order	Date	Maximum D_{st}	A_p	Solar cycle
1	13 Nov 1960	−301	280	19
2	13 Mar 1989	−599	246	22
3	1 Apr 1960	−327	241	19
4	15 Jul 1959	−429	236	19
5	18 Sep 1941		230	17
6	5 Jul 1941		220	17
7	28 Mar 1946		215	18
8	1 Mar 1941		205	17
9	6 Oct 1960	−287 (7th)	203	19
10	8 Feb 1986	−307	202	21

The storm of 13 March 1989

The storm of March 1989 (Joslyn, 1990), the largest or next to largest on record – depending on the criterion applied – had some quite remarkable effects. It was related to the largest group of sunspots to be seen on the solar disk since 1982, and the effects were not only magnetic but were also detected in the neutral atmosphere, in the ionosphere and on radio communications, as auroral displays in unusual places, and on electric-power transmission.

Magnetic effects

The storm began with a sudden commencement at 0127 UT on 13 March, and later that day the magnetic deviation at one mid-latitude station (Boulder, Colorado) amounted to more than 1300 nT. This is nearly three times the typical deviation for a K index of 9, and clearly this storm went well beyond the normal range of measurement for magnetic storms. The A_p index for 13 March was 246 (which is the second largest value recorded during the 57 years since that index was commenced in 1932), and the D_{st} index determined from equatorial ionograms reached almost -600 nT at one time. The storm continued for about two days.

The magnetic variations were large enough to have serious effects on magnetic prospecting. Whereas geophysical exploration techniques are concerned with variations of half a degree, in Alaska the magnetic declination varied by as much as 5°. Analysis of magnetometer data showed that, towards the end of 13 March, the electrojet (normally considered an auroral electrojet) was flowing south of Fredericksburg, Virginia, whose geomagnetic latitude is 49° N.

It was reported from Alaska that the flight of homing pigeons was affected.

The aurora, magnetosphere, and solar wind

Aurorae were reported at unusually low latitude in several countries. Over the western hemisphere, a red aurora was observed as far south as Florida, Mexico, and Grand Cayman Island. The key to the apparent displacement of the auroral zone, evidenced by the electrojet as well as by the luminosity, is provided by magnetic-field measurements on geosynchronous satellites. GOES-6 and GOES-7, both at $6.6R_E$, left the magnetosphere and entered the solar wind between 0700 and 0800 LT on 13 March. Making a reasonable assumption about the shape of the magnetosphere, it was deduced that the magnetopause at noon was at $4.7R_E$ instead of the usual $10R_E$. Clearly the magnetosphere was so compressed that its various internal as well as external boundaries moved to unusual places.

Unfortunately, there were no direct measurements in the solar wind during this storm. Section 2.2.3 showed that the position of the magnetopause is related to the pressure of the solar wind (Equation (2.6)); in the simplest case the distance depends on the inverse sixth root of the pressure ($2Nmv^2$). Therefore, if the subsolar magnetosphere moved in from $10R_E$ to $4.7R_e$, the solar-wind pressure would

have had to increase by a factor of 60. Some one R_E of the movement might be attributable to the southward component of the IMF (Section 2.2.3), but even so one can say with reasonable certainty that the solar-wind pressure must have increased by at least a factor of $(10/5.7)^6 = 30$ on 13 March.

The ionosphere

There was also a severe ionospheric storm on that day. The mid-latitude electron content in the night sector was unusually low immediately after the commencement. After sunrise it remained at essentially night-time values for most of the day and then it returned rapidly to a daytime value just before sunset. The equatorial ionosphere virtually disappeared, and HF radio communications were practically impossible over many circuits, particularly those involving high latitudes. However, VHF communication, which is normally restricted to line-of-sight propagation, was achieved over remarkably long distances due to high-latitude sporadic-E. An analysis of the remarkable world-wide ionospheric effects was presented by Yeh *et al.* (1992).

Satellite drag

The main effect on the neutral upper atmosphere was an increase in air density (and thus an increase in satellite drag) resulting from the heating of the atmosphere. Those whose work it is to track satellites found themselves with many more examples than usual of objects that could not immediately be identified because they were not in the places where they were expected. (In general, magnetic storms increase the rate of decay of satellites in orbit and cause them to re-enter the atmosphere sooner than predicted.)

Electric-power distribution

Perhaps the most serious consequence of this storm, however, was its effect on electric-power distribution. It is known that fluctuations in the geomagnetic field induce currents in long metallic lines (both power lines and oil pipelines). In power-distribution systems these may cause the voltage to surge, saturating transformers and tripping protective relays. The electric-power system of Québec suffered a power black-out lasting 9 h on 13 March. Users in the north-eastern USA were also affected. There was a loss of voltage on several power-distribution lines in Sweden.

Great storms may be infrequent, but they attract interest as extreme cases of the storm phenomena requiring scientific explanation, and because they may have serious effects on a number of practical activities.

2.5.6 Wave phenomena of the magnetosphere

Hydromagnetic and magnetosonic waves

In a sound wave the restoring force is due to the compressibility of the medium; in the case of a gas, its pressure. At the lowest frequencies gravity can also be significant, giving the acoustic-gravity wave (Section 1.6). In a magnetic field another restoring force comes into play, and that is the magnetic pressure across the field and the magnetic tension along the field-lines.

The situation in the magnetosphere is that the ionization may not cross field-lines, and therefore in transverse motions they must move together. Ordinary sound waves are allowed along the field-lines because the gas displacement is longitudinal, but in waves transverse to the field both the gas pressure and the magnetic pressure must be included. This combination makes possible a range of *hydromagnetic waves*.

The basic hydromagnetic wave is the *Alfvén wave*, which propagates along the magnetic field but whose displacement is transverse. The Alfvén wave is analogous to the transverse wave on a taut string, the tension being the magnetic tension (B^2/μ_0), and the mass per unit length being simply the mass density of the plasma. The speed of the transverse wave is then given by Equation (2.7):

$$v_A = \frac{B}{(\mu_0 \rho)^{1/2}},$$

where B is the magnetic flux density, μ_0 the permeability of free space, and ρ the density of the plasma in kg m^{-3}.

The Alfvén wave and the sonic wave are independent when they are traveling along the field direction, but at other angles they interact to give mixed *magnetosonic waves*. There are two such waves in general, except that perpendicular to the field there is only one, having speed $(v_A^2 + s^2)^{1/2}$, s being the speed of sound.

Micropulsations

If the sensitivity of a magnetometer is increased sufficiently, small fluctuations of the geomagnetic field with periodicites of minutes and seconds can be detected. These are *micropulsations*. They are due not to electromagnetic waves but to hydromagnetic waves in the magnetosphere, and in magnitude they are less than 10^{-4} of the total geomagnetic field. Their connection with ionospheric radio is somewhat indirect, but an introduction is in order because they comprise a significant phenomenon of the high-latitude ionosphere. In some cases they are connected with auroral activity, and there are also some diagnostic applications indicating conditions in the magnetosphere.

Micropulsations are classified according to period and duration, as in Table 2.3. The impulsive variety (Pi) occurs mainly in the evening, whereas the more

regular and persistent regular pulsations (Pc) prefer the morning and the daylight hours.

The magnetospheric origin of micropulsations is demonstrated by their similarity in magnetically conjugate regions, but a variety of mechanisms is involved in their generation. Pc pulsations are generated either at the surface of the magnetosphere or within it, and they propagate in a hydromagnetic mode. Pc1 are attributed to bunches of protons (probably) traveling back and forth between mirror points (Section 2.3.4). A resonance between protons and ion-cyclotron waves, which rotate in the same sense in the geomagnetic field, is probably involved. Pc2–5 are explained as various modes of oscillation within the magnetosphere, some propagating across and some along the field lines. The period of Pc3 and 4 may be interpreted in terms of Alfvén waves, whose speed depends on the plasma density and the magnetic field strength. The characteristic frequency changes across the plasmapause (Section 2.3.2) due to the sharp change of particle density.

The topic of micropulsations was discussed in detail by Jacobs (1970).

Instabilities

The interaction of magnetospheric waves with the particle population of the magnetosphere is a complex subject that we can no more than indicate here. For ionospheric physics its basic importance is that waves and particles may exchange energy, and that this process can become unstable. For example, trapped electrons generate whistler-mode waves in the VLF band (Section 3.4.7), which, under the right conditions, may then interact with the population of trapped electrons, scattering them into the loss cone (Section 2.3.4). This is a mechanism that, thereby, may cause the spontaneous precipitation of trapped electrons into the atmosphere. There is a large literature on *wave–particle interactions* in the magnetosphere. The interested reader might like to start with Lyons and Williams (1984), Chapter 5.

One purely ionospheric instability is the *two-stream*, or *Farley–Buneman*, instability, which produces electrostatic waves in the E-region electrojet when

Table 2.3. *Micropulsations*

Continuous and regular		Irregular	
Type	Period (s)	Type	Period (s)
Pc1	0.2–5	Pi1	1–40
Pc2	5–10	Pi2	40–150
Pc3	10–45		
Pc4	45–150		
Pc5	150–600		

streams of ions and electrons differ in velocity by more than a critical amount. This is the mechanism causing the irregularities in ionization that make auroral radar possible. That topic is discussed in Sections 3.5.1, 4.2.2, and 6.5.5.

The *Kelvin–Helmholz instability* is a commonplace phenonenon, being the cause of waves on the surface of a pond on a windy day. It works because any projection above the level surface alters the air flow in such a way as to increase the perturbation – a simple case of positive feedback. The magnetosphere also has interfaces, the most obvious being that with the solar wind at the magnetopause, and Kelvin–Helmholtz waves and vortices may be produced there.

A slightly less superficial introduction to magnetospheric waves can be found in Hargreaves (1992), Chapter 9.

2.6 Ionization by energetic particles

The main source of ionization in the upper atmosphere is solar radiation in the X-ray and EUV bands. There is, however, another source of a quite different kind, namely energetic particles. Although they are not entirely absent from middle latitudes, these are much more important at high latitudes, where they may at times become the main source of ionization. As we shall see in later chapters, two very significant sources at high latitude are electrons associated with the aurora, and protons (plus some α-particles) emitted from the Sun during some solar flares.

2.6.1 Electrons

Various methods have been used to calculate the rate of ion production by a stream of energetic electrons arriving from some source above the atmosphere. The most generally useful one relies on laboratory measurements of the range of electrons in air. An electron loses energy to the neutral gas particles with which it collides, and the rate of loss obviously depends on the number of gas particles encountered. Thus, in a uniform atmosphere the distance traveled varies in inverse proportion to the gas pressure. The unit of range (r_0) is therefore [pressure] × [distance]: atm cm, for example. The energy goes into exciting and ionizing the neutral particles. In this instance we are interested in the ionization.

An energetic particle entering the atmosphere from above travels into a medium of increasing density, and the altitude, h_p, to which it penetrates is such that the product of pressure and distance, integrated above h_p, is equal to the range r_0. Obviously, this particle will ionize only above height h_p, and the total number of ion–electron pairs produced will depend on $E/\Delta E$, where E is the initial energy of the particle and ΔE is the energy required for each ionization (generally taken as 34 or 35 eV).

The third fact which has to be taken into account is that the rate of energy loss, and therefore of ion production, along the path is a function of the particle's

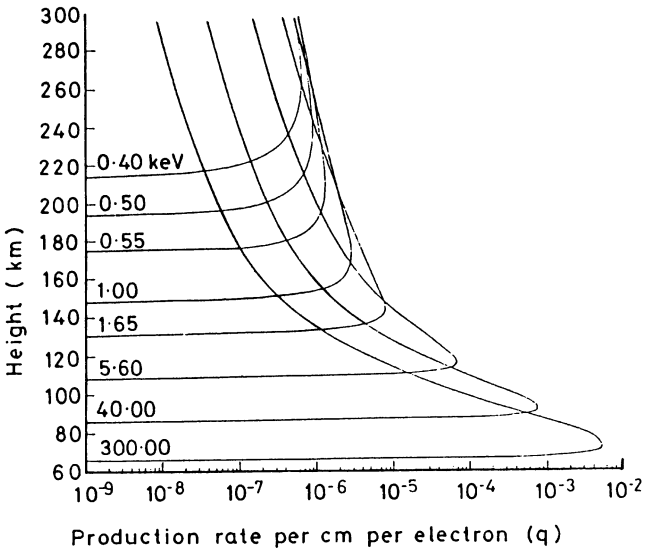


Figure 2.26. Production rates due to monoenergetic electrons of various initial energies. (After M. H. Rees, *Planet. Space Sci.* **11**, 1209, Copyright 1963, with permission from Elsevier Science.)

velocity. To quantify this, laboratory measurements are again brought to bear in the form of an "efficiency", which is a function of the atmospheric depth at a point along the track divided by that at the point where the particle eventually stops; i.e. s/s_p. (The atmospheric depth is the total mass of gas in a column of unit cross-section along the path of the particle.) The efficiency is normalized to unity over the whole path (from $s/s_p = 0$ to $s/s_p = 1$), and it comes to a maximum at $s/s_p = 0.4$ for a monoenergetic electron beam traveling directly along the magnetic field. If the electrons arrive over a range of angles, as would be the case in nature, some particles travel in spiral paths and so cover a greater distance; in this case the efficiency peaks at a smaller value of s/s_p such as 0.1 or 0.2. These factors clearly influence the distribution of ion production in the atmosphere, and the height of the maximum production rate in particular.

Figure 2.26 shows calculated production rates in a model atmosphere due to monoenergetic electrons of various initial energies. Note that the production rate peaks at lower altitude and the distribution is narrower for higher initial enery. To get the effect of a more realistic spectrum, the production rate must be integrated over energy.

2.6.2 *Bremsstrahlung* X-rays

When energetic electrons collide with neutral gas particles a small amount of their energy is converted to X-rays through the *Bremsstrahlung* process – literally "braking radiation" – as they are rapidly decelerated. The X-rays penetrate deeper into the atmosphere than do the primary electrons and may be observed by balloon-borne detectors at heights of 30–40 km. Some X-rays are also scattered back out of the atmosphere and may be detected on satellites.

The computation of the *Bremsstrahlung* X-ray flux is fairly complicated, since an electron of energy E can produce photons of energy E or less, and the X-rays are emitted over a wide range of angles. Inverting an observed X-ray spectrum to give the spectrum of the primary electrons is even more difficult. The usual practice is to draw on a set of computations giving the *Bremsstrahlung* due to single electrons of specified energy, but even so it is usually neccessary to assume a form for the spectrum (e.g. exponential).

When the X-rays are stopped by the atmosphere they create ionization at that level. The ionization rate due to *Bremsstrahlung* is several factors of ten smaller than that due to the primary electrons higher up, but, at the height concerned, possibly 50 km or below, it is the major source of ionization at times of auroral electron precipitation. Table 2.4 compares the heights and maximum production rates due to direct and *Bremsstrahlung* ionization for several initial electron energies.

Figure 2.27 illustrates the relative altitude and magnitude of direct and X-ray ionization (actually the rate of deposition of energy) due to a spectrum of incident electrons with characteristic energy 10 keV.

2.6.3 Protons

Significant ionization may also be caused by energetic protons, especially at high latitudes during *polar-cap events*, which are due to fluxes of protons released from the Sun at the time of a solar flare. A lesser flux of α-particles will generally arrive simultaneously. These particles, which are significantly more energetic than the auroral electrons discussed above, lose energy in colliding with the atmospheric gas and leave ionized trails. The gas concerned is principally that of the mesosphere, whose composition is essentially like that of the troposphere, and therefore the rate of energy loss is well known from laboratory measurements. A graph showing the rate of energy loss against the distance traveled is called a *Bragg curve*. In the energy range of interest to us the loss rate increases as the proton slows down, and, over the range 10–200 MeV, the loss rate is almost inversely proportional to the energy, a typical value being 0.8 MeV per meter of path in air at standard temperature and

Table 2.4. *Direct and X-ray ion production*

	Height of maximum production (km)		Maximum production rate (ion pairs cm^{-3} per electron)	
E (keV)	Direct	X-ray	Direct	X-ray
3	126	88	2.5×10^{-5}	5.9×10^{-10}
10	108	70	1.4×10^{-4}	1.3×10^{-8}
30	94	48	5.6×10^{-4}	2.3×10^{-7}
100	84	37	1.9×10^{-3}	1.3×10^{-5}

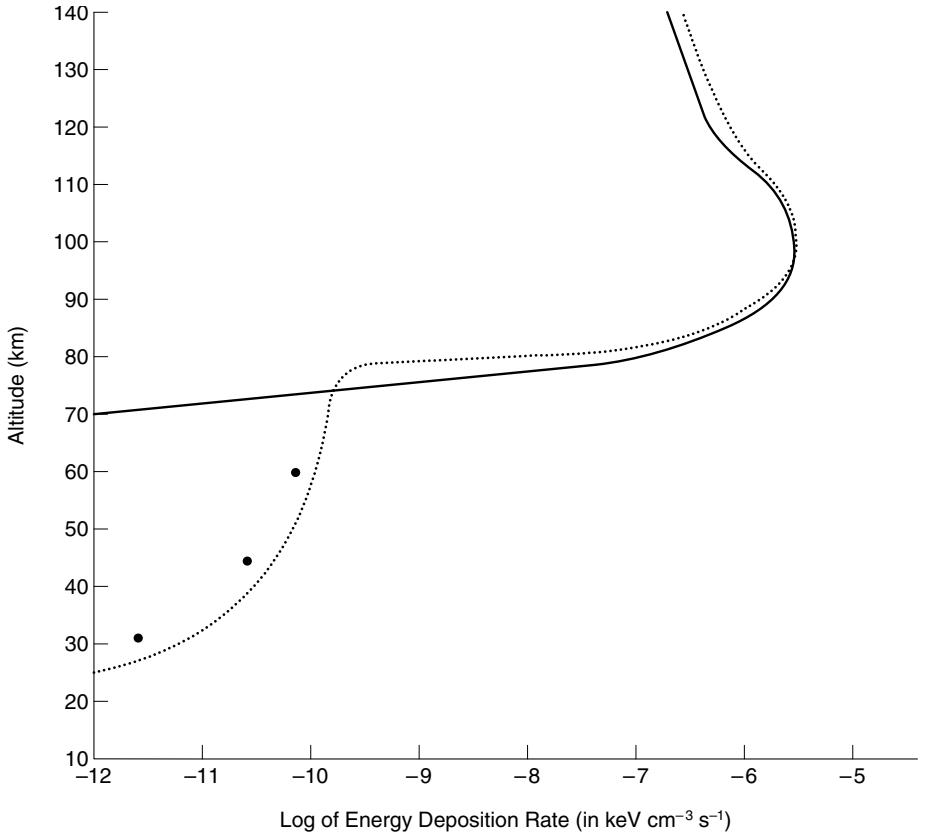


Figure 2.27. A comparison of direct and X-ray energy-deposition rates from an incoming electron flux with an exponential spectrum of characteristic energy 10 keV. The solid line (J. G. Luhmann, *J. Atmos. Terr. Phys.* **38**, 605, 1976) shows the deposition by electron impact. The dashed line (M. J. Berger *et al.*, *J. Atmos. Terr. Phys.* **36**, 591, 1974) and the circles (J. G. Luhmann, *J. Atmos. Terr. Phys.* **39**, 595, 1977) include the energy deposited by *Bremsstrahlung* X-rays. (After Luhmann 1977, Copyright, with permission from Elsevier Science.)

pressure when the energy is 100 MeV. The energy may be assumed to be used entirely in creating ion–electron pairs, each requiring about 35 eV.

The nature of the Bragg curve, combined with the density distribution of the atmosphere, means that the ionization due to a proton entering the atmosphere from space is very concentrated towards the end of the path. A vertically incident 50 MeV proton, for example, loses half its energy over the last 2.5 km of the path and the last 10% over only 100 m. One consequence is that the penetration level does not depend strongly on the angle of incidence except near 90°. Production-rate profiles for protons of various initial energies are given in Figure 2.28. Note the low altitudes which may be reached by the more energetic particles. For a spectrum of proton energies the total effect would be calculated by appropriate summing over these curves at each height.

There is a similar procedure for dealing with the ionization by α-particles.

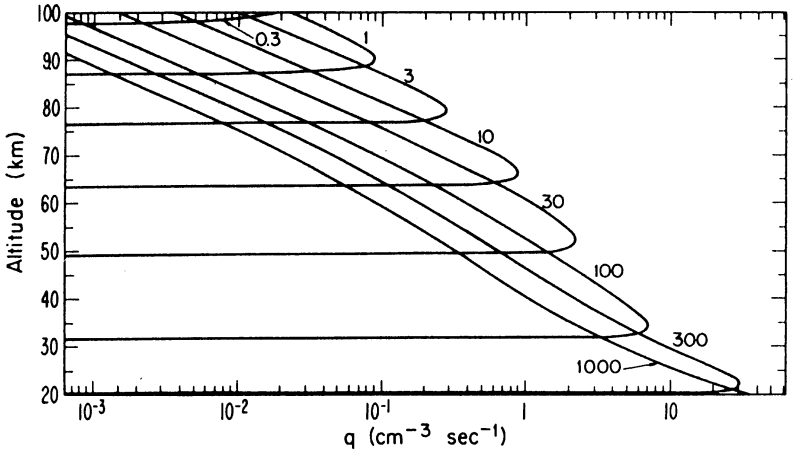


Figure 2.28. Production rates due to incident monoenergetic protons. The initial energies are given in MeV, and in each case the flux is 1 proton cm^{-2} s^{-1} sr^{-1}. (G. C. Reid, *Fundamentals Cosmic Phys.* **1**, 167, 1974. Copyright OPA (Overseas Publishers Association) NV, with the permission of Gordon and Breach Publishers.)

2.7 References and bibliography

2.2 The magnetosphere

Burlaga, L. F. (1982) Magnetic fields in the interplanetary medium. *Solar System Plasmas and Fields* (eds. J. Lemaire and M. J. Rycroft), p. 51. Pergamon Press, Oxford.

Carpenter, D. L. and Smith, R. L. (1964) Whistler measurements of electron density in the magnetosphere, *Rev. Geophys.* **2**, 415.

Gosling, J. T. (1972) Predicting the solar wind speed. *Solar Activity Observations and Predictions* (eds. P. S. McIntosh and M. Dryer), p. 231. MIT Press, Cambridge Massachusetts.

Holzer, T. E. (1979) The solar wind and related astrophysical phenomena. *Solar System Plasma Physics* (eds. C. F. Kennel, L. J. Lanzerotti and E. N. Parker), Vol. I, p. 103. Elsevier Science Publishers, Amsterdam.

Pertinec, S. M. and Russell, C. T. (1993) External and internal influences on the size of the dayside terrestrial magnetosphere. *Geophys. Res. Lett.* **20**, 339.

Raitt, W. J. and Schunk, R. W. (1983) Composition and characteristics of the polar wind. *Energetic Ion Composition in the Earth's Magnetosphere* (ed. R. G. Johnson), p. 99. Terra Scientific Publishing Co., Tokyo.

Rich, F. J. and Hairston, M. (1994) Large-scale convection patterns observed by DMSP. *J. Geophys. Res.* **99**, 3827.

Smith, E. J., Tsurutani, B. T. and Rosenberg, R. L. (1978) Observations of the interplanetary sector structure up to heliographic latitudes of 16°: Pioneer 11. *J. Geophys. Res.* **83**, 717.

Vasyliunas, V. M. (1983) Large-scale morphology of the magnetosphere. *Solar–Terrestrial Physics* (eds. R. L. Carovillano and J. M. Forbes), p. 243. Reidel, Dordrecht.

Wilcox, J. M. and Ness, N. F. (1965) Quasi-stationary corotating structure in the interplanetary medium. *J. Geophys. Res.* **70**, 5793.

2.3 Particles in the magnetosphere

Burch, J. L. (1977) The magnetosphere. *Upper Atmosphere and Magnetosphere*, p. 42. National Academy of Sciences, Washington DC.

Carpenter, D. L. and Anderson, R. R. (1992) An ISEE/whistler model of equatorial electron density in the magnetosphere. *J. Geophys. Res.* **97**, 1097.

Chappell, C. R., Harris, K. K. and Sharp, G. W. (1970) A study of the influence of magnetic activity on the location of the plasmapause as measured by OGO-5. *J. Geophys. Res.* **75**, 50.

Helliwell, R. A. (1976) *Whistlers and Related Ionospheric Phenomena*. Stanford University Press, Stanford, California.

Iijima, T. and Potemra, T. A. (1978) Large-scale characteristics of field-aligned currents associated with substorms. *J. Geophys. Res.* **83**, 599.

Potemra, T. A. (1983) Magnetospheric currents. *Johns Hopkins APL Tech. Digest* **4**, 276.

Ratcliffe, J. A. (1972) *An Introduction to the Ionosphere and Magnetosphere*. Cambridge University Press, Cambridge.

Van Allen, J. A. (1959) The geomagnetically trapped corpuscular radiation. *J. Geophys. Res.* **64**, 1683.

Walt, M. (1994) *Introduction to Geomagnetically Trapped Radiation*. Cambridge University Press, Cambridge.

Williams, D. J. (1985) Dynamics of the Earth's ring current: theory and observation. *Space Sci. Rev.* **42**, 375.

2.4 Dynamics of the magnetosphere

Dungey, J. W. (1963) The structure of the exosphere or adventures in velocity space. In *Geophysics, The Earth's Environment* (eds. C. De Witt, J. Hieblot, and A. Lebeau). Gordon and Breach, New York.

Levy, R. H., Petschek, H. E., and Siscoe, G. L. (1964) Aerodynamic aspects of the magnetospheric flow. *Am. Inst. Aeronaut. Astronaut. J.* **2**, 2065.

Nishida, A. (1966) Formation of plasmapause, or magnetospheric plasma knee, by the combined action of magnetospheric convection and plasma escape from the tail. *J. Geophys. Res*. **71**, 5669.

Petschek, H. E. (1964) Magnetic field annihilation. *The Physics of Solar Flares* (ed. W. N. Hess), Report SP-50, p. 425. NASA, Washington DC.

Ratcliffe, J. A. (1972) *An Introduction to the Ionosphere and Magnetosphere*. Cambridge University Press, Cambridge.

Russell, C. T. (1972) The configuration of the magnetosphere. *Critical Problems of Magnetospheric Physics*, p.1. IUCSTP Secretariat, National Academy of Science, Washington D.C.

2.5 Magnetic storms

Akasofu, S.-I. (1968) *Polar and Magnetospheric Substorms*. Reidel, Dordrecht.

Davis, T. N. and Sugiura, M. (1966) Auroral electrojet activity index AE and its universal time variations. *J. Geophys. Res*. **71**, 785.

Hargreaves, J. K. (1992) *The Solar–Terrestrial Environment*. Cambridge University Press, Cambridge.

Jacobs, J. A. (1970) *Geomagnetic Micropulsations*. Springer-Verlag, Berlin.

Joslyn, J. A. (1990) Case study of the great magnetic storm of 13 March 1989. *Astrodynamics* (eds. Thornton, Proulx, Prussing and Hoots).

Lyons, L. R. and Williams, D. J. (1984) *Quantitative Aspects of Magnetospheric Physics*. Reidel, Dordrecht.

Maynaud, P. N. (1980) *Derivation, Meaning and Use of Geomagnetic Indices*. American Geophysical Union, New York.

Sugiura, M. and Chapman, S. (1960) The average morphology of geomagnetic storms with sudden commencement. *Abhandl. Akad. Wiss. Göttingen Math.-Phys. Kl.* Special Issue **4**, 53.

Yeh, K. C., Lin, K. H. and Conkright, R. O. (1992) The global behavior of the March 1989 ionospheric storm. *Can. J. Phys*. **70**, 532.

2.6 Ionization by energetic particles

Berger, M. J., Seltzer, S. M. and Maeda, K. (1974) Some new results on electron transport in the atmosphere. *J. Atmos. Terr. Phys*. **36**, 591.

Luhmann, J. G. (1976) Auroral electron spectra in the atmosphere. *J. Atmos. Terr. Phys*. **38**, 605.

Luhmann, J. G. (1977) Auroral bremsstrahlung spectra in the atmosphere. *J. Atmos. Terr. Phys*. **39**, 595.

Rees, M. H. (1963) Auroral ionization and excitation by incident energetic electrons. *Planet. Space Sci.* **11**, 1209.

Reid, G. C. (1974) Polar-cap absorption – observations and theory. *Fundamentals Cosmic Phys.* **1**, 167.

General reading on the topics of Chapter 2

Books

Akasofu, S.-I. and Chapman, S. (1972) *Solar–Terrestrial Physics*. Oxford University Press, Oxford.

Baumjohann, W. and Treumann, R. A. (1996) *Basic Space Plasma Physics*. Imperial College Press, London.

Carovillano, R. L. and Forbes, J. M. (eds.) (1983) *Solar–Terrestrial Physics*. Reidel, Dordrecht.

Carovillano, R. L., McClay, J. F. and Radoski, H. R. (1968) *Physics of the Magnetosphere*. Springer-Verlag, New York.

Hess, W. N. (1968) *The Radiation Belt and Magnetosphere*. Blaisdell, Waltham, Massachusetts.

Hess, W. N. and Mead, G. D. (eds.) (1968) *Introduction to Space Science*. Gordon and Breach, New York.

Hundhausen, A. J. (1972) *Coronal Expansion and Solar Wind*. Springer-Verlag, New York.

Jacobs, J. A. (1970) *Geomagnetic Micropulsations*. Springer-Verlag, New York.

Jursa, A. S. (ed.). (1985) *Handbook of Geophysics and the Space Environment*. Air Force Geophysics Laboratory, US Air Force, National Technical Information Service, Springfield, Virginia.

Kamide, Y. (1988) *Electrodynamic Processes in the Earth's Ionosphere and Magnetosphere*. Kyoto Sangyo University Press, Kyoto.

Le Galley, D. P. and Rosen, A. (eds) (1964) *Space Physics*. Wiley, New York.

Lyons, L. R. and Williams, D. J. (1984) *Quantitative Aspects of Magnetospheric Physics*. Reidel, Dordrecht.

Nishida, A. (ed.) (1982) *Magnetospheric Plasma Physics*. Reidel, Dordrecht.

Parks, G. K. (1991) *Physics of Space Plasmas*. Addison-Wesley Publishing Co., Redwood City, California.

Roederer, J. G. (1974) *Dynamics of Geomagnetically Trapped Radiation*. Springer-Verlag, Berlin.

Schulz, M. and Lanzerotti, L. J. (1974) *Particle Diffusion in the Radiation Belts*. Springer-Verlag, New York.

Treumann, R. A. and Baumjohann, W. (1997) *Advanced Space Plasma Physics*. Imperial College Press, London.

Conference reports

Akasofu, S.-I. (ed.) (1980) *Dynamics of the Magnetosphere*. Reidel, Dordrecht.

Beynon, W. J. G., Boyd, R. L. F., Cowley, S. W. H. and Rycroft, M. J. (1989) *The Magnetosphere, the High-Latitude Ionosphere, and their Interactions*. The Royal Society, London.

Johnson, R. G. (ed.) (1983) *Energetic Ion Composition in the Earth's Magnetosphere*. Terra Scientific Publishing Co., Tokyo.

Kamide, Y. and Slavin, J. A. (eds.) (1986) *Solar Wind–Magnetosphere Coupling*. Terra Scientific Publishing Co., Tokyo.

King, J. W. and Newman, W. S. (eds.) (1967) *Solar–Terrestrial Physics*. Academic Press, London.

Lemaire, J. F., Heynderickx, D. and Baker, D. N. (eds.) (1996) *Radiation Belts: Models and Standards*. American Geophysical Union, Washington DC.

McCormac, B. M. (ed.) (1966) *Radiation Trapped in the Earth's Magnetic Field*. Reidel, Dordrecht.

McCormac, B. M. (ed.) (1968) *Earth's Particles and Fields*. Reinhold, New York.

McCormac, B. M. (ed.) (1970) *Particles and Fields in the Magnetosphere*. Reidel, Dordrecht.

McCormac, B. M. (ed.) (1972) *Earth's Magnetospheric Processes*. Reidel, Dordrecht.

McCormac, B. M. (ed.) (1974) *Magnetospheric Physics*. Reidel, Dordrecht.

McCormac, B. M. (ed.) (1976) *Magnetospheric Particles and Fields*. Reidel, Dordrecht.

Olsen, W. P. (ed.) (1979) *Quantitative Modeling of Magnetopheric Processes*. American Geophysical Union, Washington DC.

Potemra, T. A. (ed.) (1984) *Magnetospheric Currents*. American Geophysical Union, Washington DC.

Song, P., Sonnerup, B. U. O. and Thomsen, M. F. (eds.) (1995) *Physics of the Magnetopause*. American Geophysical Union, Washington DC.

Tsurutani, B. T., Gonzalez, W. D., Kamide Y. and Arballo, J. K. (1997) *Magnetic Storms*. American Geophysical Union, Washington DC.