Cambridge Books Online

http://ebooks.cambridge.org/

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 6 - The aurora, the substorm, and the E region pp. 285-336

Chapter DOI: http://dx.doi.org/10.1017/CBO9780511535758.008

Cambridge University Press

Chapter 6

The aurora, the substorm, and the E region

6.1 Introduction

By "aurora" people usually mean the emission of light from the upper atmosphere, but in fact there are numerous related phenomena, each being a direct or indirect consequence of energetic particles entering the atmosphere from the magnetosphere. They include

(a) luminous aurora;
(b) radar aurora, by which is meant the reflection of radio signals from ionization in the auroral region;
(c) auroral radio absorption, the absorption of radio waves in the auroral ionization;
(d) auroral X-rays, which are generated by the incoming particles and may be detected on high-altitude balloons;
(e) magnetic disturbances, due to enhanced electric currents flowing in the auroral ionization, which may be detected by magnetometers;
(f) electromagnetic emissions in the very-low- and ultra-low-frequency bands, which are generated in the magnetosphere by wave–particle interactions (Section 2.5.6), and which then propagate to the ground where they may be detected with a radio receiver or a sensitive magnetometer.

Arising as they do from a common cause, the auroral phenomena display several common properties.

(1) They all exhibit a general relationship with solar activity, though often there is no specific association with any obvious solar event. From the

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

1930s the term *M region* was used to signify a hypothetical and unseen solar region causing aurora and magnetic storms, and this served as a unifying hypothesis for some 40 years. It is now well appreciated, of course, that the unseen agent is the solar wind.

(2) They are essentially zonal in occurrence, their occurrence and intensity coming to a maximum some 10°–25° from the magnetic poles. This property is treated in Section 6.2.

(3) All the auroral phenomena exhibit *substorm* behavior. They are greatly enhanced during bursts of activity lasting perhaps 30–60 min, which are separated by quieter intervals of several hours. It is now clear that the substorm is caused by processes in the magnetosphere. This aspect is discussed in Section 6.4.

The auroral luminosity originates within the ionospheric E region. The particles which excite the emission of light also create additional ionization and thereby enhance the electron density. This in turn increases the ionospheric current at those heights, which has further consequences. The behavior of the auroral E region is therefore closely related to that of the aurora. The high-latitude E region is considered in Section 6.5.

6.2 Occurrence zones

6.2.1 The auroral zone and the auroral oval

In general the auroral phenomena are highly structured in both space and time, with essentially zonal patterns of occurrence. The classical picture of the occurrence of aurorae (Figure 6.1) shows a zone centered about 23° from the geomagnetic pole (i.e. about 67° geomagnetic latitude) and about 10° wide in latitude. The isochasms show the occurrence of discrete aurorae, which is 100% at the maximum and falls off both to the equatorward and to the poleward sides. ("100%" here means that some aurora was seen every clear night, not that it was visible all the time.) This pattern, which is a geographic distribution, was first defined by Vestine (1944) and is based on reports of visual sightings of the aurora over several decades.

However, in 1963 Y. I. Feldstein, using all-sky camera data from the International Geophysical Year of 1957–1958, pointed out that at a fixed time the locus of the aurora is not circular but oval (Figure 6.2). The maximum is near 67° latitude at midnight, but this increases to about 77° (the latitude of the cusp – Section 2.2.5) at noon. The *auroral oval*, as it is generally known, is widest at midnight and narrowest at noon. It is essentially fixed with respect to the Sun, and the classical auroral zone is the locus of the midnight sector of the oval as the Earth rotates underneath it. The auroral oval is one of the important boundaries of geospace. In relation to the structure of the magnetosphere it is generally considered

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

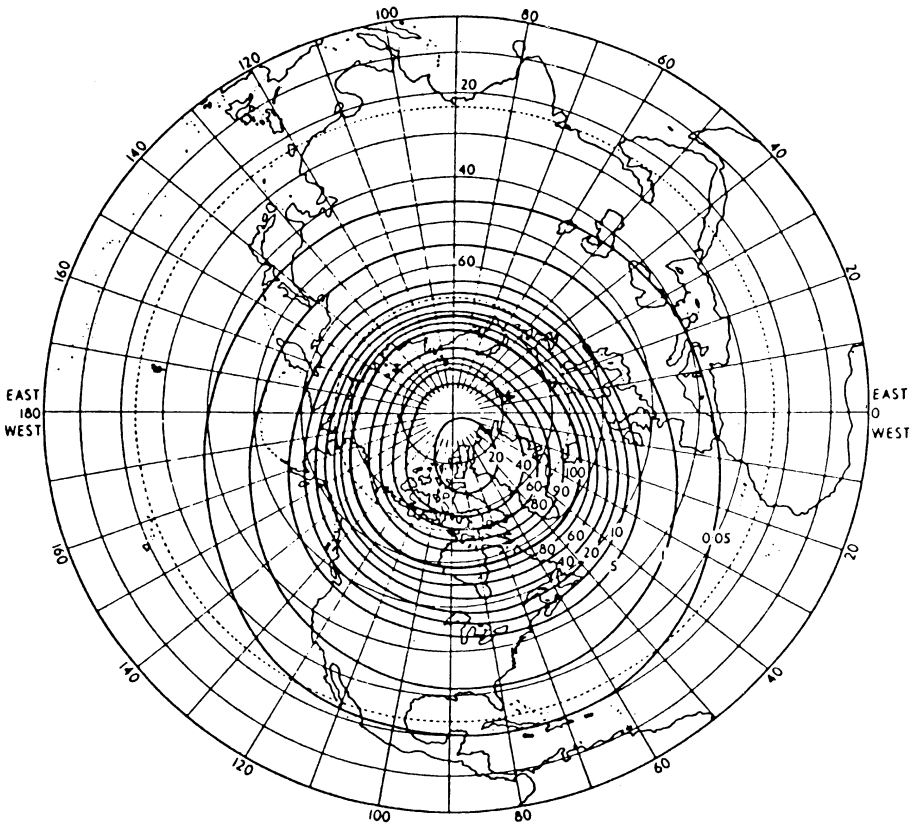


Figure 6.1. The northern auroral zone, showing the percentage of good observing nights when aurorae may be seen. (After E. H. Vestine, *Terr. Magn. Atmos. Electricity*, **49**, 77, 1944, copyright by the American Geophysical Union.)

to mark the division between open and closed field-lines. The regions poleward of the ovals (one in each hemisphere) are generally taken to be the polar caps in which the magnetic field-lines connect to the IMF and circulate under the influence of the solar wind (Section 2.4.1).

Although it was originally just a statistical concept, later work, both ground-based (Feldstein and Starkov, 1967) and using photography from space (Akasofu, 1974; Frank and Craven, 1988), has shown that the oval exists virtually continuously as a permanent ring of light around the magnetic pole, and also as a ring of particle precipitation (Fuller-Rowell and Evans, 1987; Hardy *et al.*, 1985). In the pictures from space the general form of the oval is clearly observed (Figure 6.3) as a continuous band of light around the pole that is nearly always present, though its intensity varies greatly from time to time.

The latest auroral pictures from space are adding much detail and have shown that the oval form is not by any means the whole story. The detailed spatial distribution varies considerably from time to time. Sometimes an arc is seen to extend across the polar cap, connecting the day and night sides of the oval – a configuration called the *θ-aurora*. Sometimes the morning side of the oval is quiet while the

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

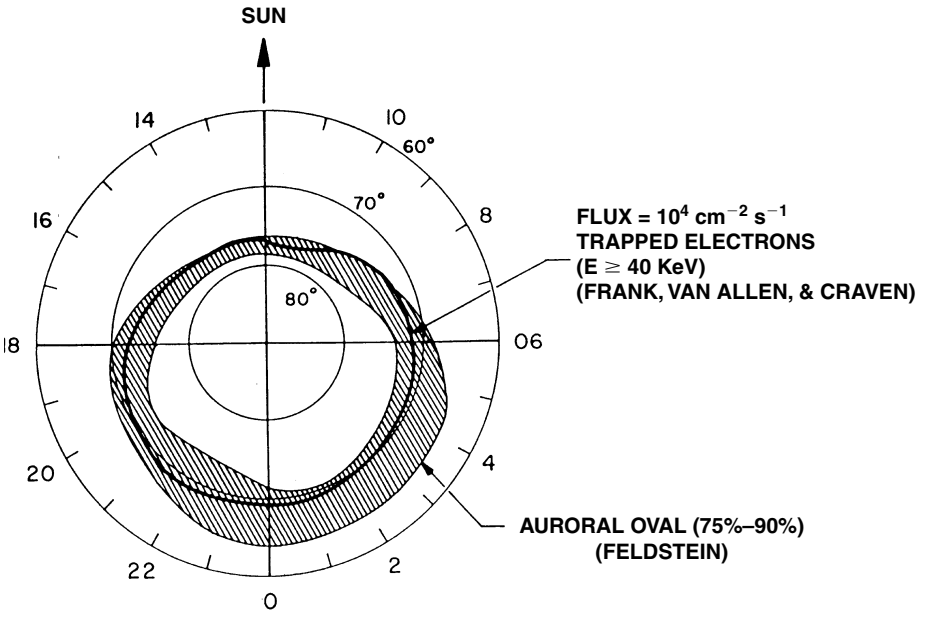


Figure 6.2. The auroral oval in relation to the 40-keV trapping boundary. (S.-I. Akasofu, *Polar and Magnetospheric Substorms*, Reidel, 1968, with kind permission from Kluwer Academic Publishers.)

evening is active, and sometimes the morning side is the more active. The behavior of more localized brightenings within the oval can also be observed from space. Some examples are shown in Figure 6.4. The variety of configurations and dynamics emphasizes the complexity of the auroral distribution and suggests that present classifications are incomplete.

6.2.2 Models of the oval

Without doubt the auroral oval is a special region of the ionosphere. That being so, it is often convenient to refer observed phenomena to the location (or probable location) of the oval at the time of the observation, and thus it is helpful to have models giving the typical position of the oval under stated conditions. Figure 6.5(a) indicates, for typical conditions, the geographic location of the oval every 2 h of the UT day. It is usual to quantify the level of disturbance by using one of the magnetic activity indices (see Section 2.5.4), Q being a popular one for this purpose since it is derived at 15-min intervals. Figure 6.5(b) gives the locations of the oval for several levels of Q taken from a set of diagrams developed by Whalen (1970). (K_p being a more common index, the following relations may be used to obtain the appropriate value of Q when one is using Whalen's results: $Q=8$ if $K_p \geqslant 6$; $Q=K_p+2$ if $1<K_p<6$; and $Q=3K_p$ if $K_p \leqslant 1$.) Since the oval is closer to the magnetic pole at noon than it is at midnight, it is quite possible for an observer on the Earth to be poleward of the oval at midnight and equatorward of it at noon.

Distributions based not on luminosity but on measurements of particles of

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

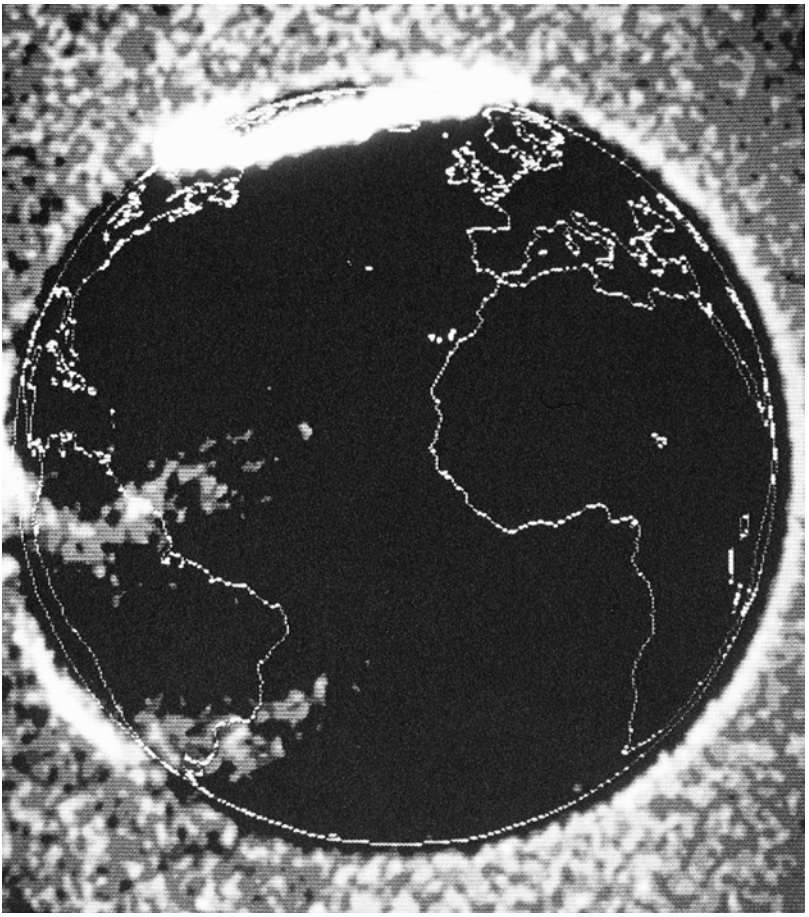


Figure 6.3. The auroral oval from space, observed in the ultra-violet (118–165 nm) by the Dynamics Explorer I spacecraft on 16 February 1982. The aurora is plainly visible around the northern pole. Airglow bands north and south of the equator, dayglow above the morning (right) limb of the Earth, and resonant Lyman-α scattering in the protonosphere are also to be seen. (L. A. Frank and J. D. Craven, University of Iowa, private communication.)

energy 30 or 50 eV to 20 keV on DMSP satellites (Hardy *et al.*, 1985; 1989) show zones of electron and ion influx in terms of magnetic latitude and local time, at levels of geomagnetic activity quantified by K_p. Figure 6.6 shows the distribution of electrons for $K_p = 3$. Figure 6.6(a) is very like the auroral oval, being offset from the magnetic pole towards midnight. The particles forming the dayside maximum are relatively soft (i.e. of low energy).

Using data from the same source, Meng and Makita (1986) defined the boundaries of the precipitation zones for "low-energy" ($<$500 eV) and "high-energy" ($>$500 eV) electrons, for magnetically quiet (AE$\leqslant$150 nT) and disturbed (AE$>$400 nT) conditions, and for the evening and morning sectors – see Table 6.1. The criterion for the boundary was a flux of 10^7 electrons cm^{-2} s^{-1} steradian^{-1}. The transition latitude, where the fluxes of low- and high-energy particles were equal, was also noted.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

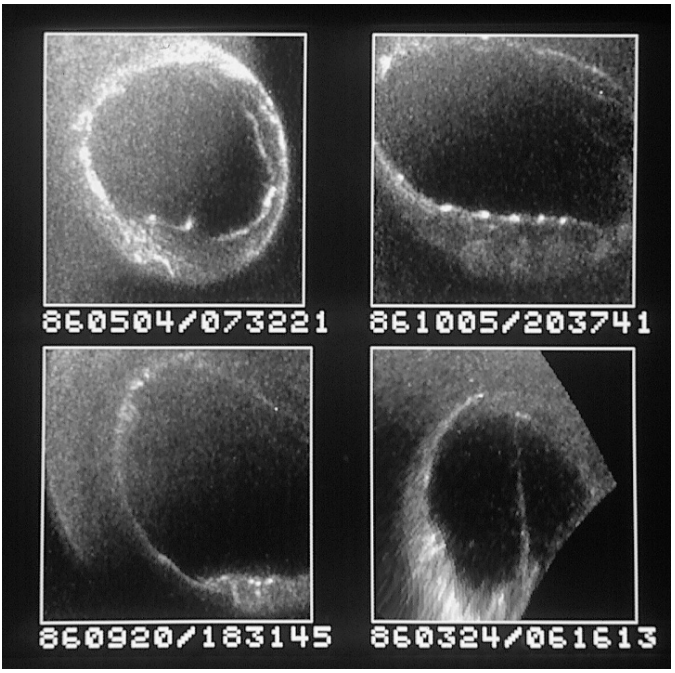


Figure 6.4. Images of the northern auroral region observed by the Viking satellite. The camera had a 20° by 25° field of view and responded to UV of 134–180 nm, mainly emissions from nitrogen. Each exposure lasted 1.2 s. The top left-hand image shows the whole auroral oval including the day side. The one below it shows a substorm in the midnight sector, with activity also around noon and faint arcs in the morning. The top right-hand image is from the last stage of a substorm, when regularly spaced bright spots, lasting 1–5 min, may appear along the poleward edge of the oval near midnight. The fourth image shows a sun-aligned arc extending across the polar region from midnight (at the bottom) to noon. (Pictures and commentary from G. Enno, private communication. The Viking project was managed by the Swedish Space Corporation for the Swedish Board for Space Activites. The UV imager was a project of the National Research Council of Canada, and was operated by the Institute for Space Research, University of Calgary, with support from the Natural Sciences and Engineering Research Council of Canada.)

It is generally agreed that the oval expands equatorward, to lower latitudes, as magnetic activity increases. The region of low-energy precipitation becomes narrower and the high-energy region broadens. According to Chubb and Hicks (1970), the equatorward boundary of the luminous oval moves about 1.7° equatorward per unit of K_p on the day side of the Earth, and 1.3° on the night side; it moves by 1°–3° of latitude in individual substorms. (See also Section 6.4.2.) The oval also varies with the IMF, increasing in size by about 0.5° for each one γ increase in the southward component of the IMF. According to Meng (1984), the polar cap can be as small as 12° side to side under quiet conditions and as large as 50° when it is disturbed, which is not inconsistent with the results in Table 6.1.

Gussenhoven *et al.* (1983) express the variation of the equatorward boundary of the oval in terms of the K_p index, as

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

$$L = L_0 + aK_p, \tag{6.1}$$

where L_0 and a depend on the local magnetic time (MLT) as in Table 6.2.

Figure 6.7 illustrates the position of the oval at three levels of disturbance, and its magnetic latitude and thickness against K_p.

The foregoing results may be expected to apply also to those propagation phenomena which are typical of the auroral oval.

6.3 The auroral phenomena

6.3.1 The luminous aurora

The luminous aurora is a well-known phenomenon of the high-latitude regions, and in fact the most readily observed consequence of the dynamic magnetosphere. Although it is only in the present century that there has been any kind of understanding of the aurora, it must surely rank amongst the oldest of the known geophysical phenomena. There are accounts of lights in the night sky going back to Greek and Roman times, when they were frequently given a mystical or prophetic interpretation. The term *aurora borealis* dates from 1621, and the southern lights, observed by Captain James Cook in 1773, were subsequently called the *aurora australis*. Detailed reports of auroral displays date from 1716 and the first written work devoted entirely to the polar aurora was published in France in 1733.

The first proof that the auroral light is a consequence of the excitation of atmospheric gas by energetic particles did not come until the early 1950s, and it was not until 1958, when rockets were fired into an aurora, that energetic electrons were identified as the primary source. Where those electrons come from and how they are energized are questions that have not yet been answered in full, but their magnetospheric origin is beyond doubt and much has been learned about them in recent years.

6.3.2 The distribution and intensity of the luminous aurora

Auroral investigations before about 1950 tended to fall into one of two areas. Morphological studies were intended to map the occurrence of the aurora in space and time and to determine the details of the fine structure of individual auroral forms. Auroral spectroscopy was virtually a separate discipline, concerned with the emitted light, in particular with its spectrum and its origin in photochemical processes – a topic having strong affinities with airglow.

The luminous aurora is highly structured and dynamic. Some features are as thin as 100 m, and the time changes can be as rapid as 10 s^{-1}. The basic recording intrument is the *all-sky camera* which was first used during the 1950s and is particularly valuable for surveying the occurrence of aurorae. It uses a convex

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

Figure 6.5a (1)

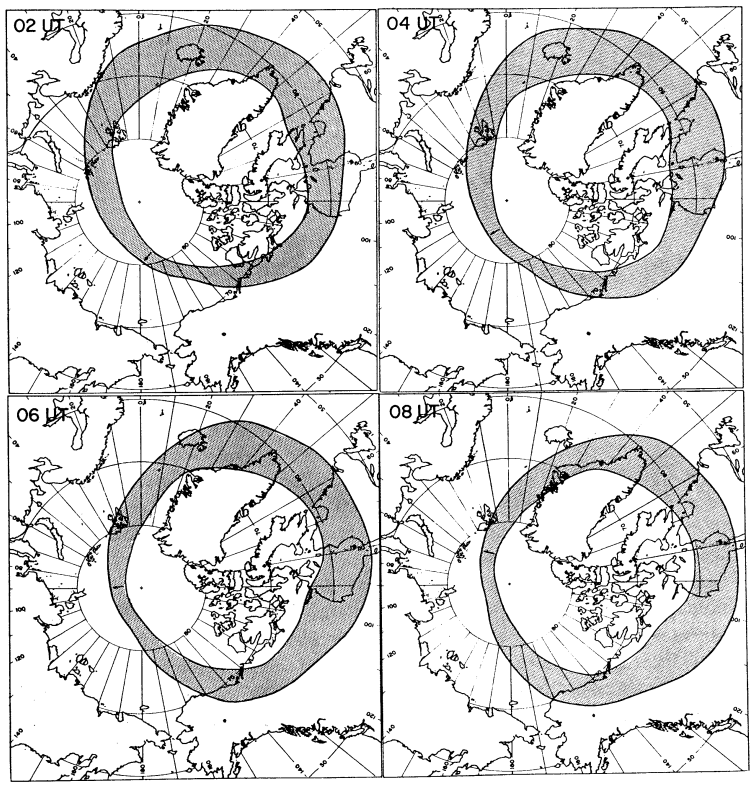


Figure 6.5a (2)

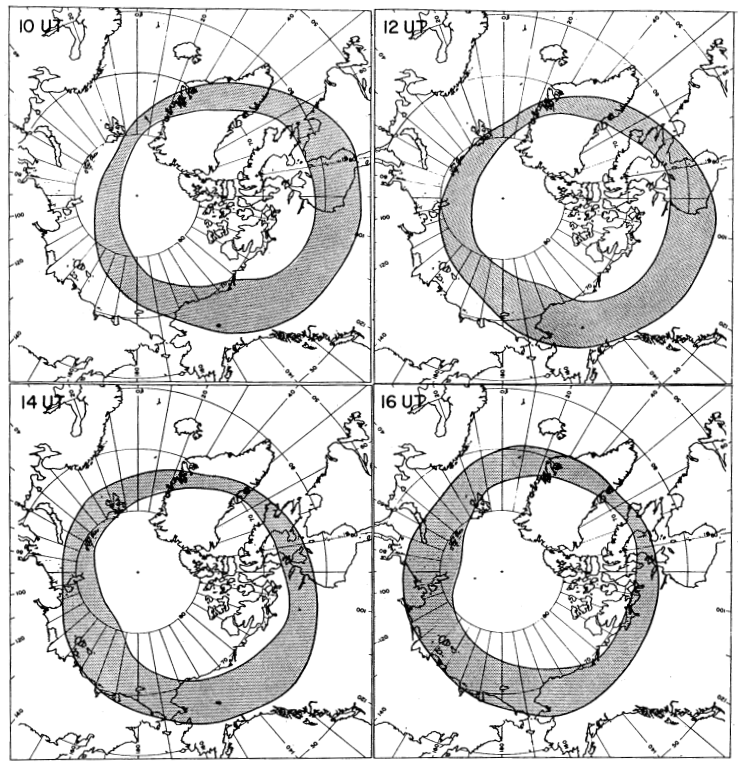


Figure 6.5a (3)

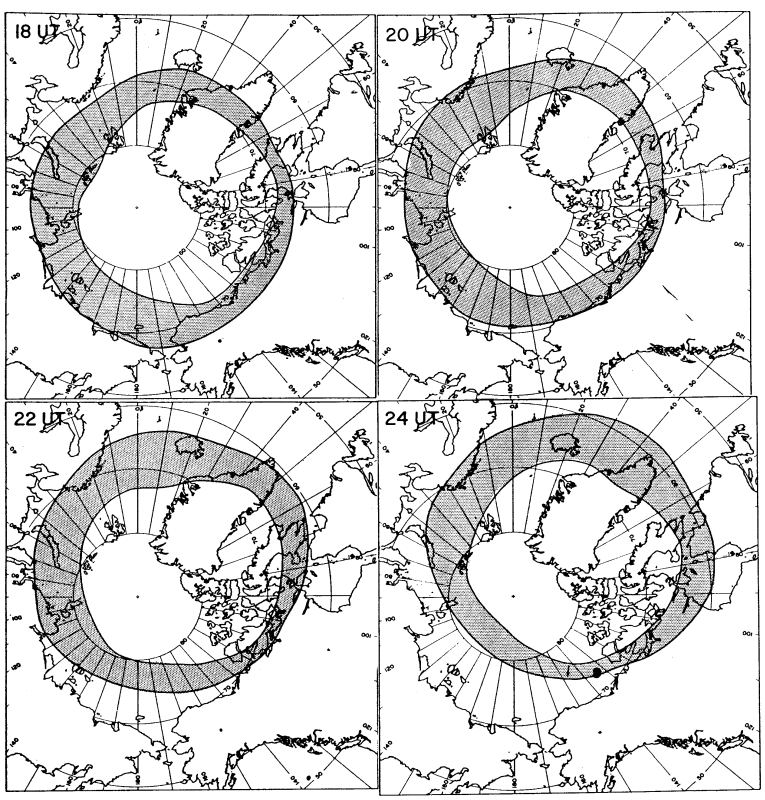
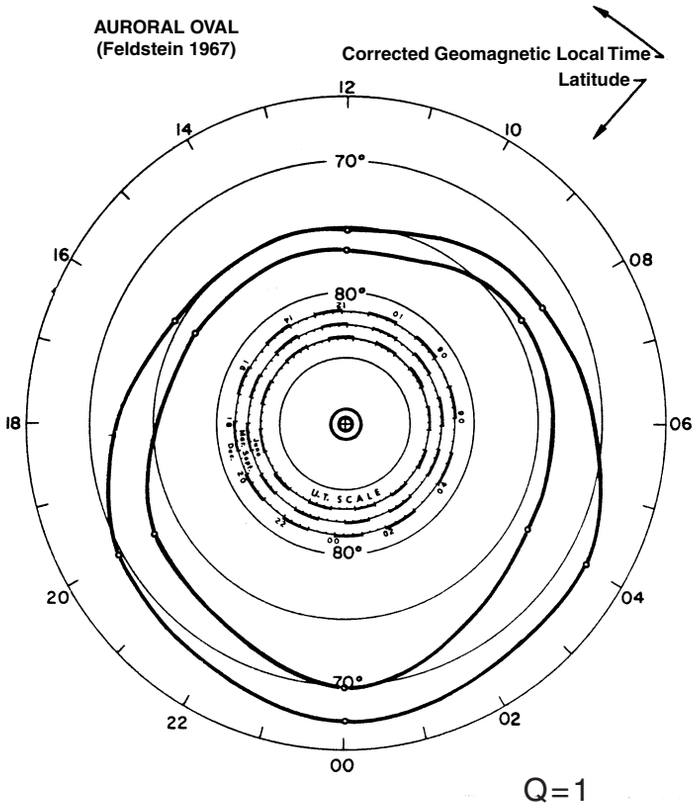


Figure 6.5. Representations of the auroral oval. (a) The geographic position of the auroral oval under typical conditions for each 2 h (UT) of the day. (S.-I. Akasofu, *Polar and Magnetospheric Substorms*, Reidel, 1968, with kind permission from Kluwer Academic Publishers.) (b) The position of the oval in geomagnetic coordinates at disturbance levels $Q = 1$, 3, 5 and 7. (J. A. Whalen. Report AFCRL-70-0422, 1970.)

Cambridge Books Online © Cambridge University Press, 2009



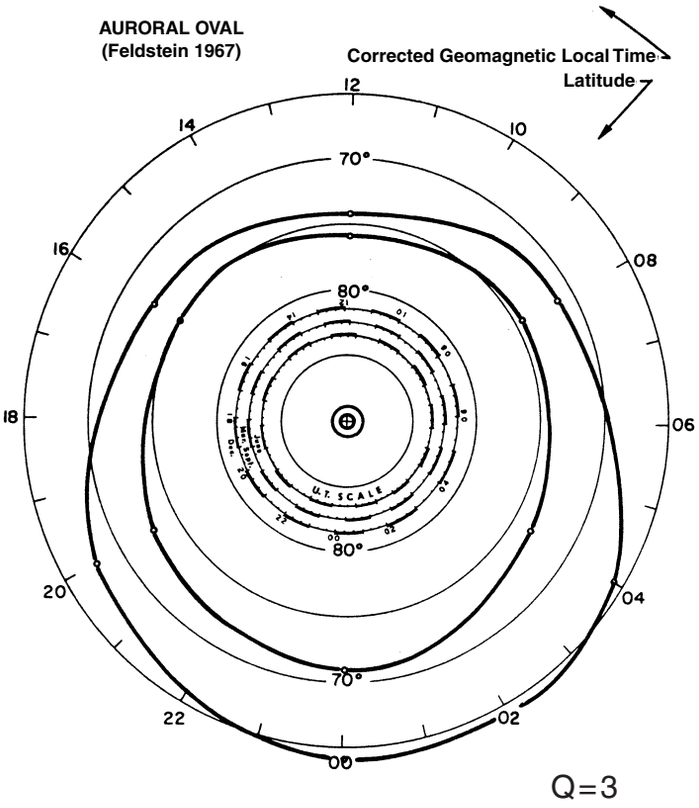
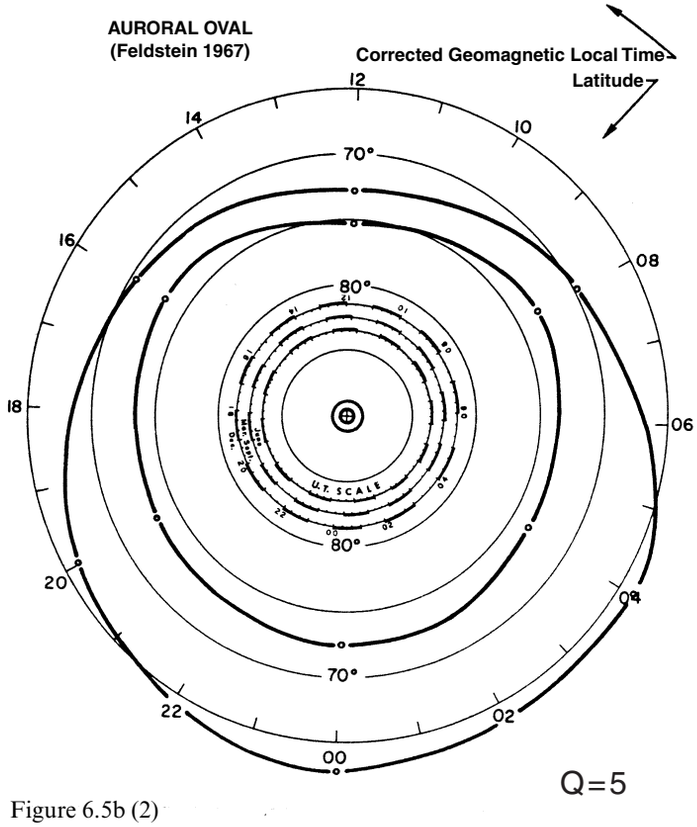


Figure 6.5b (1)



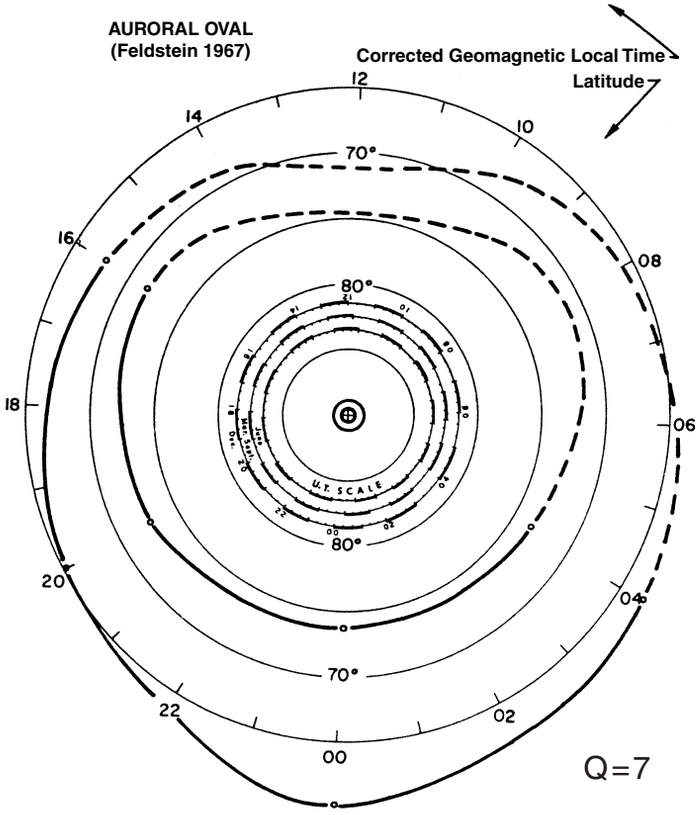


Figure 6.5b (2)

Cambridge Books Online © Cambridge University Press, 2009

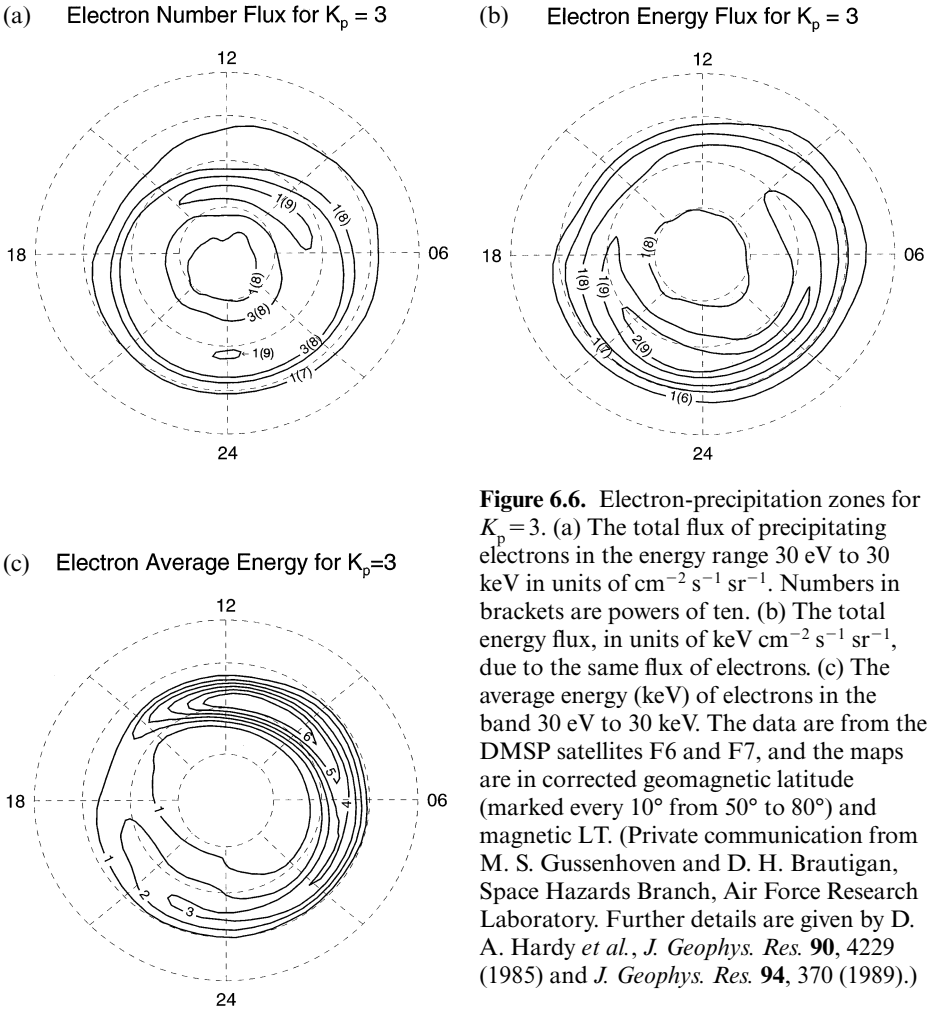


Figure 6.6. Electron-precipitation zones for $K_p = 3$. (a) The total flux of precipitating electrons in the energy range 30 eV to 30 keV in units of $cm^{-2}\ s^{-1}\ sr^{-1}$. Numbers in brackets are powers of ten. (b) The total energy flux, in units of keV $cm^{-2}\ s^{-1}\ sr^{-1}$, due to the same flux of electrons. (c) The average energy (keV) of electrons in the band 30 eV to 30 keV. The data are from the DMSP satellites F6 and F7, and the maps are in corrected geomagnetic latitude (marked every 10° from 50° to 80°) and magnetic LT. (Private communication from M. S. Gussenhoven and D. H. Brautigan, Space Hazards Branch, Air Force Research Laboratory. Further details are given by D. A. Hardy *et al.*, *J. Geophys. Res.* **90**, 4229 (1985) and *J. Geophys. Res.* **94**, 370 (1989).)

mirror to obtain a picture of the night sky from horizon to horizon, and it would typically be operated automatically at regular intervals on every clear night during the winter viewing season.

There is a classification of auroral structure based on its general appearence, as in Table 6.3. When structure is present the height of the luminosity may be determined by triangulation. Between 1911 and 1943, C. Störmer made 12000 height determinations with spaced cameras and found that the lower borders of auroral forms were usually at heights of 100–110 km (Figure 6.8(a)). In some of the forms the luminosity is concentrated into a band only 10–20 km deep and the lower edge, in particular, can be quite sharp. The brightness of a discrete arc typically falls off by a factor of ten within a few kilometers below the maximum, and by a further factor of ten only 1 or 2 km below that. The vertical distribution of auroral luminosity is illustrated in Figure 6.8(b) for several types of aurora.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

Table 6.1. *The magnetic latitude of the poleward boundary of "low"-energy, and the equatorward boundary of "high"-energy electron precipitation (after Meng and Makita, 1986)*

	Quiet conditions		Disturbed conditions	
	Evening	Morning	Evening	Morning
Poleward boundary (low energy)	80°–82°	80°–82°	73°–75°	76°–77°
High-to-low-energy transition	73°–75°	73°–75°	70°–72°	70°–72°
Equatorward boundary (high energy)	69°–71°	67°–69°	64°–66°	64°–66°

Table 6.2. *Values of L_0 and a for Equation (6.1)*

MLT (h)	L_0	a
00–01	66.1	−1.99
01–02	65.1	−1.55
04–05	67.7	−1.48
05–06	67.8	−1.87
06–07	68.2	−1.90
07–08	68.9	−1.91
08–09	69.3	−1.87
09–10	69.5	−1.69
10–11	69.5	−1.41
11–12	70.1	−1.25
12–13	69.4	−0.84
15–16	70.9	−0.81
16–17	71.6	−1.28
17–18	71.1	−1.31
18–19	71.2	−1.74
19–20	70.4	−1.83
20–21	69.4	−1.89
21–22	68.6	−1.86
22–23	67.9	−1.78
23–24	67.8	−2.07

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

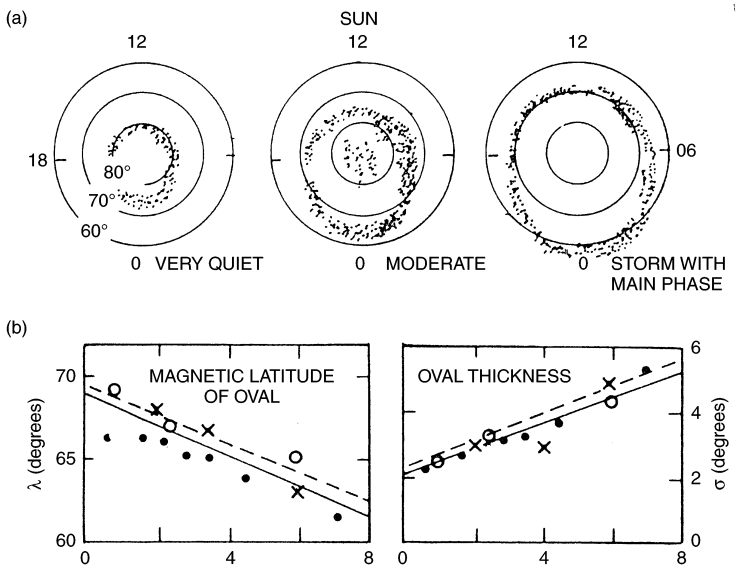


Figure 6.7. (a) Positions of the auroral oval under three levels of activity. (b) The magnetic latitude and thickness of the oval as functions of Kp. (J. M. Goodman, *HF Communications*. Van Nostrand Reinhold, 1992.)

Table 6.3. *Classification of auroral forms*

Forms without ray structure

Homogeneous ray structure: a luminous arch stretching across the sky in a magnetically east–west direction; the lower edge is sharper than the upper, and there is no perceptible ray structure

Homogeneous band: somewhat like an arc but less uniform, and generally exhibiting motions along its length: the band may be twisted into horseshoe bends

Pulsating arc: part or all of the arc pulsates

Diffuse surface: an amorphous glow without distinct boundary, or isolated patches resembling clouds

Pulsating surface: a diffuse surface that pulsates

Feeble glow: auroral light seen near the horizon, so that the actual form is not observed

Forms with ray structure

Rayed arc: a homogeneous arc broken up into vertical striations

Rayed band: a band made up of numerous vertical striations

Drapery: a band made up of long rays, giving the appearance of a curtain; the curtain may be folded

Rays: ray-like structures, appearing singly or in bundles separated from other forms

Corona: a rayed aurora seen near the magnetic zenith, giving the appearance of a fan or a dome with the rays converging on one point

Flaming aurora: waves of light moving rapidly upward over an auroral form

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

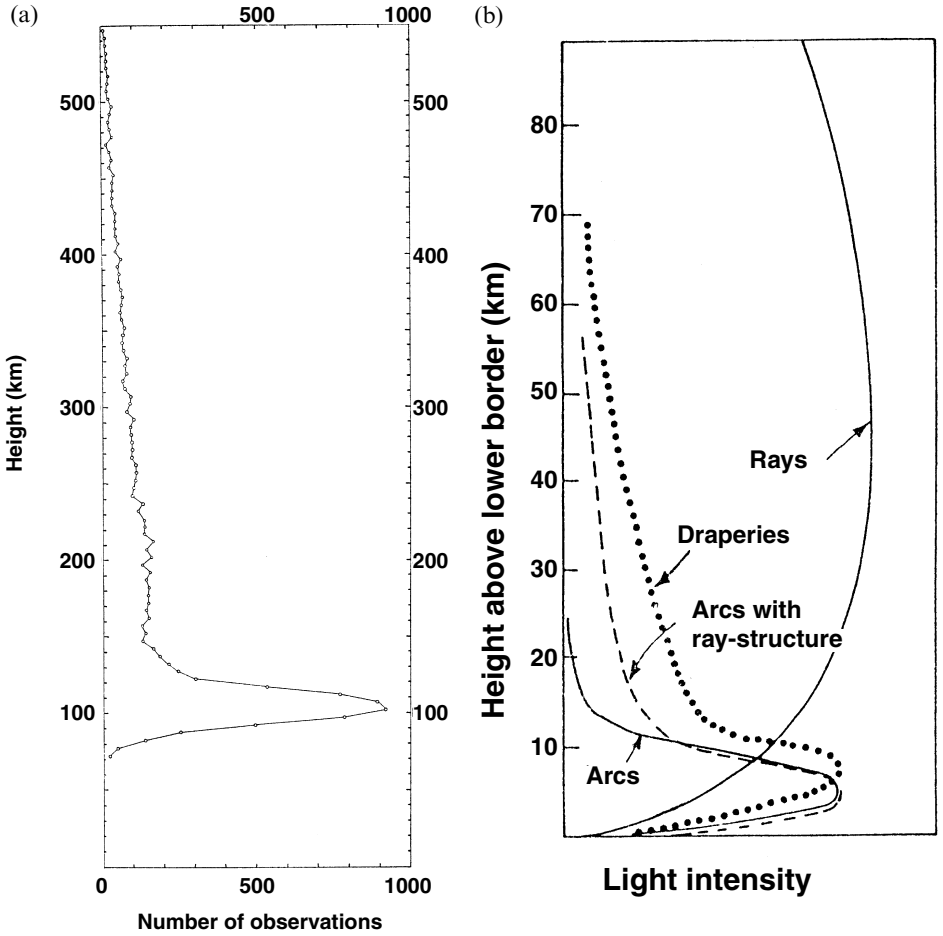


Figure 6.8. Observations of auroral luminosity. (a) The distribution of 12330 height measurements made by Störmer and colleagues. The vast majority lie between 90 and 150 km. (C. Störmer, *The Polar Aurora*. Oxford University Press, 1955. By kind permission of Oxford University Press.) (b) Profiles of auroral luminosity along various forms. (After L. Harang, *The Aurorae*. Wiley, 1951.)

The intensity of an aurora as seen from the ground is measured in units of the Rayleigh, named in honour of R. J. Strutt (the fourth Baron Rayleigh), who was a notable amateur scientist of his time and the leading pioneer of airglow studies. The unit (R) is defined as

$$1\ \mathrm{R} = 10^6\ \text{photons cm}^{-2}\ \text{s}^{-1}. \tag{6.2}$$

It is a measure of the height-integrated rate of emission, as would be observed by an instrument on the ground looking vertically upward. A more general classification of brightness uses a scale of I–IV, as in Table 6.4. This also shows the

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

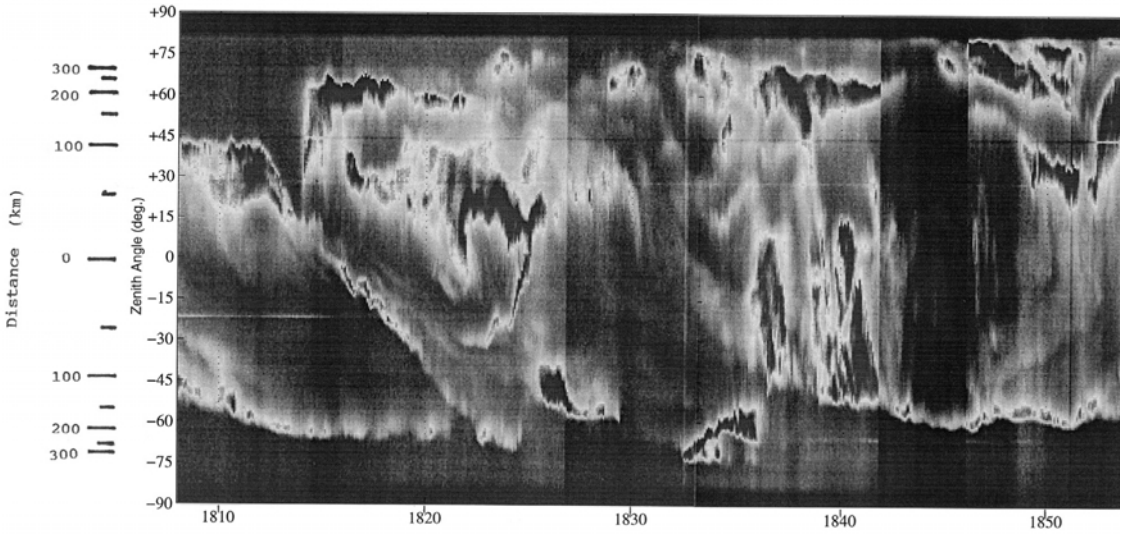


Figure 6.9. A keogram from auroral TV in Scandinavia. North (poleward) is at the top. This example shows the main features of auroral activity during a 2-h period on 18 February 1993, giving some idea of the complexity of the aurora on an active day ($K_p = 4$). Note the dominance of equatorward movements. There are also several poleward

standard a visual observer might use for comparison, the equivalent in kilorayleighs, and the approximate rate of deposition of energy into the atmosphere.

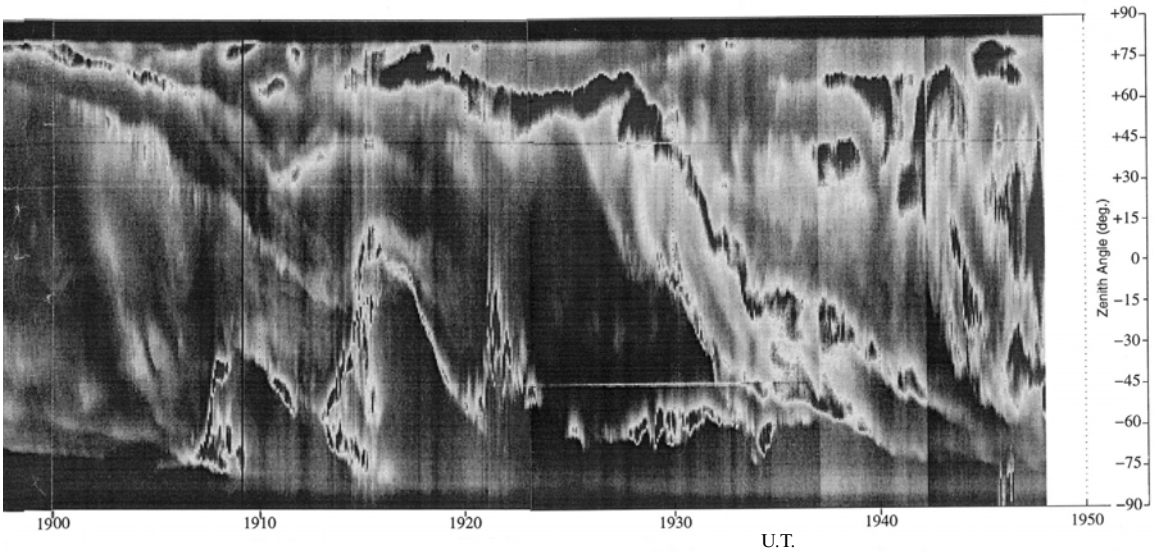
A photometer is needed for exact intensity measurements, and this can either be pointed in a fixed direction, for example to the zenith, or scanned across the sky to record the spatial distribution of intensity as well. A diagram of the latitudinal variations with time is sometimes called a *keogram*. Neither scanning photometers nor cameras are sufficiently sensitive to record the most rapid fluctuations in the auroral emissions, but TV techniques, both monochrome and color, are more sensitive and have been applied very successfully to dynamic auroral photography in recent years. In addition to their scientific value, some of these auroral "videos" are possessed of no little esthetic interest (particularly if they are set to music). Figure 6.9 shows an example of a keogram composed from TV data.

One significant distinction that should be made is that between *discrete* and *diffuse* (or *mantle*) aurora. All the earlier studies concentrated on the discrete

Table 6.4. *Intensity classification of the aurora*

Intensity	Equivalent to	Kilorayleighs	Energy deposition (erg cm^{-2} s^{-1})
I	Milky Way	1	3
II	Thin moonlit cirrus	10	30
III	Moonlit cumulus	100	300
IV	Full moonlight	1000	3000

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016



expansions, either substorm onsets or pseudo-break-ups (Section 6.4.2). The distance scale assumes that the emission comes from a height of 110 km. A distance of 600 km is equivalent to approximately 5.5° of magnetic latitude. (Data from P. N. Smith, Space Physics Group, University of Sussex, via the Auroral TV Database.)

auroral forms (Table 6.3) which are the more readily observed against the background light of the night sky because of their fine and dynamic structure. However, as was demonstrated in the early 1960s, the aurora may also take the form of a diffuse glow. This contributes at least as much total light as the discrete forms, though it is more difficult to observe from the ground because of its low intensity per unit area. The night-time discrete and diffuse aurorae are thought to map along the geomagnetic field into different regions of the magnetotail (Section 2.2.6 and Figure 2.6). The diffuse aurora is generally associated with the central part of the plasma sheet, and the discrete forms, which tend to appear poleward of the diffuse aurora, are thought to map onto the edge of the plasma sheet or to an X-type neutral line (Section 2.4.2 and Figure 2.20).

Downward-looking satellites, by virtue of their ability to observe a large part or even the whole of the auroral oval at the same time, and which avoid the problem of poor seeing conditions which so often affects the ground-based techniques, have provided much new information about the distribution of the luminous emissions. The diffuse aurora tends to dominate in these pictures, but discrete forms are also seen within the diffuse glow or poleward of it; they are not seen on the equatorward side, however.

When the IMF is northward, luminous arcs extending for thousands of kilometers and aligned towards the Sun are observed in the polar caps. They are not bright (emitting only tens of rayleighs, against thousands for a normal aurora) but they can be detected with modern equipment and at that low intensity are observed about half the time. It appears, therefore, that they are almost always

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

present when the IMF is northward. It is believed that these arcs are on closed field-lines and that they may be magnetically conjugate (i.e. they occur simultaneously at opposite ends of field-lines in northern and southern hemispheres). The Sun-aligned arcs are associated with velocity shears in the polar-cap convection (Section 5.1.2 and Figure 5.5). (The θ-aurora, mentioned above, is also associated with a velocity shear, but it is much brighter and also much rarer than the common Sun-aligned arcs. It is not at present clear whether it is a different phenomenon.)

6.3.3 Auroral spectroscopy

Aurorae and airglow have similar causes, both being the emission of quanta of radiation from common atmospheric gases, particularly O and N_2. In the first case the excitation is by energetic particles entering the upper atmosphere from the magnetosphere, and in the second by electromagnetic radiation from the Sun. The emission lines represent transitions between energy states of the emitting species, but these may be complex and the task of interpreting the auroral spectrum was far from trivial. In spectroscopists' terms the lines are in general "forbidden," which means in practice that they are generated by transitions that are relatively improbable.

Most aurorae are too faint to be seen in color by the naked eye, but a bright aurora appears green or red, the colours being due to atomic-oxygen lines at 557.7 nm (*the green line*) and 630.0 nm (*the red line*), respectively. The 391.4-nm line from ionized molecular nitrogen (N_2^+) is also present in the violet. Some aurorae have red lower borders, and, when this occurs, the red light is due to emissions from molecular oxygen. Such aurorae result from unusually energetic particles that penetrate deeper into the atmosphere. The important group of emissions from atomic oxygen and the transitions which cause them are illustrated in Figure 6.10.

Some of the UV emissions, particularly the O emissions near 130 nm, have proved particularly valuable for mapping the aurora from space vehicles because, at those wavelengths, the aurora may be seen in sunlight. In addition to their obvious applications to the detection and mapping of the aurora, some of the emission lines can be exploited to provide information helpful to other branches of upper-atmospheric science. The intensities of the emissions from N_2^+ at 427.8 and 391.4 nm are proportional to the rate of ionization by the incoming electrons. The neutral-air wind in the thermosphere may be determined by measuring the Doppler shift of the 630-nm line of oxygen.

6.3.4 Ionospheric effects

The auroral phenomena are all associated with the precipitation of energetic electrons into the atmosphere. Although the best known of them, the luminosity is actually a byproduct of ionization by energetic particles and of the subsequent recombination processes. Other phenomena, more directly related to the enhanced electron density of the auroral region, are of greater direct concern in

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

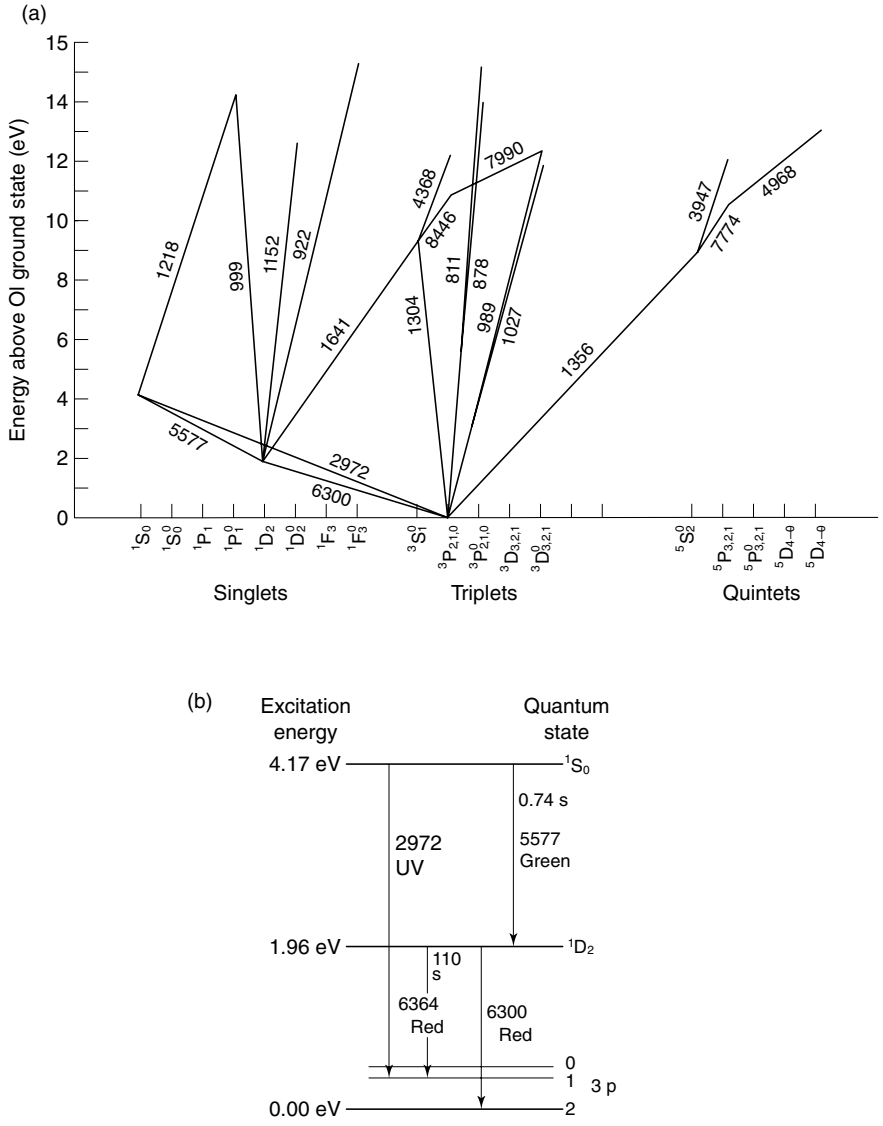


Figure 6.10. Energy levels and transitions in atomic oxygen. (a) Transitions that have been observed in airglow or aurorae. (M. H. Rees, *Physics and Chemistry of the Upper Atmosphere*. Cambridge University Press, 1989.) (b) Details of the most important lines. (After S. J. Bauer, *Physics of Planetary Ionospheres*. Springer-Verlag, 1973, copyright notice of Springer-Verlag.) In each case the unit of wavelength is the ångstrom unit.

radio propagation. They are briefly reviewed here, and some will be treated in detail in later chapters. Figure 6.11 illustrates the connections in a schematic, and admittedly simplistic, manner.

The E region

In the E region, for example, it is not unusual for the electron density to be increased to several times 10^{12} m^{-3} by electron precipitation. Electron densities of

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

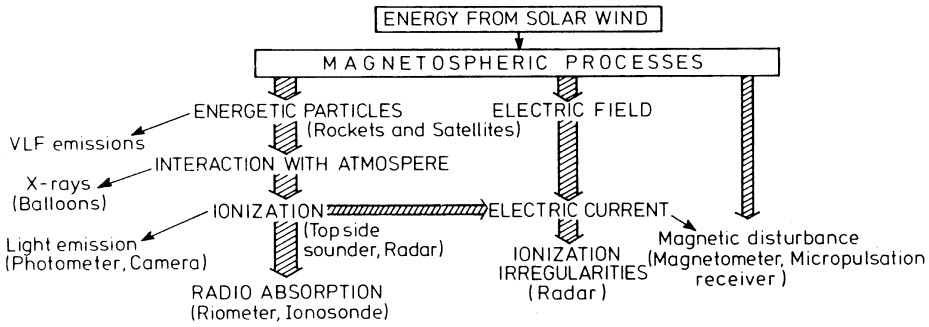


Figure 6.11. Some links between auroral phenomena. Techniques are shown in brackets. (J. K. Hargreaves, *Proc. Inst. Electr. Electron. Engineers*. **57**, 1348, © 1969 IEEE.)

this magnitude may reflect vertically incident waves of radio frequency up to 20 MHz (Section 3.4.2, Equation (3.64)), and those of higher frequency if they are obliquely incident (Section 3.4.3, Equation (3.73)). The ionization may therefore be detected by radar as a total reflection if the frequency is not too high. If the observing geometry is suitable, echoes may also be received at higher frequencies, and these echoes come from electron-density irregularities that are produced by instabilities arising in the auroral electrojet (Section 2.5.3). Since the irregularities tend to be field-aligned, the echo intensity is aspect-sensitive and the best observing geometry is when the radar lies in the plane normal to the magnetic field-lines. The radar aurora is described in detail in Section 6.5.5.

The D region

The more energetic electrons penetrate into the D region (See Figure 2.26), and the ionization they create there acts to absorb radio waves by an amount depending on their frequency. The effect is usually monitored with a Riometer (Section 4.2.4), which typically operates in the range 30–50 MHz, at which frequencies the absorption rarely exceeds 10 dB; but the effect will generally be considerably greater in the HF band. (The absorption varies approximately with the inverse square of the frequency – Section 3.4.4, Equation (3.95)). The properties of auroral radio absorption are detailed in Section 7.2. Figures 7.23 and 7.24 illustrate E- and D-region electron-density profiles observed by incoherent-scatter radar during electron-precipitation events.

X-rays

Auroral X-rays are generated by the *Bremsstrahlung* process outlined in Section 2.6.2. They have no direct influence on radio propagation, but, because of their greater penetrating power, they produce ionization at a lower altitude than do their parent electrons. The incidence and morphology of auroral X-rays are in many ways similar to those of auroral radio absorption, both being due to the harder end of the electron spectrum.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

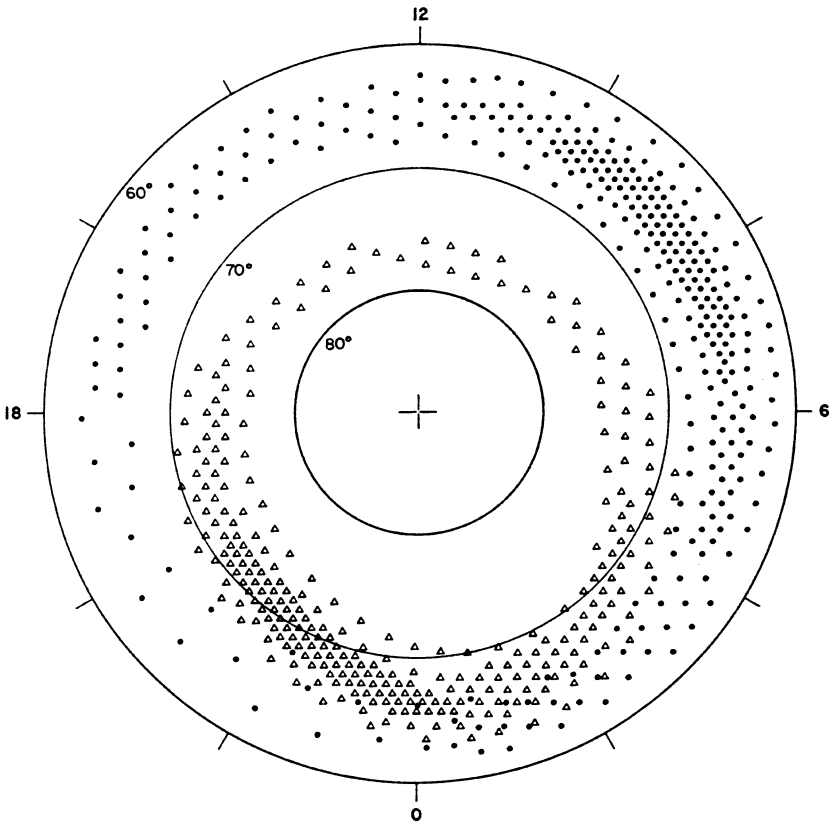


Figure 6.12. The two zones of auroral particle precipitation in the northern hemisphere. The density of symbols indicates the average flux, and the coordinates are geomagnetic latitude and time. (Reprinted from T. R. Hartz and N. M. Brice, *Planet. Space Sci.* **15**, 301, copyright 1967, with permission from Elsevier Science.)

Magnetic effects

Magnetic bays (Section 2.5.3) are essentially a phenomenon of the auroral zone, though, as a magnetic perturbation, they are also detected by magnetometers a considerable distance away. The bays are primarily due to the ionospheric current which flows in enhancements of the E-region electron density. VLF and ULF emissions also increase when the auroral zone is active. They have various causes, some involving wave–particle interactions, but they are basically magnetospheric in origin and are not a factor in radio propagation.

6.3.5 The outer precipitation zone

Hartz and Brice (1967) generalized the definition of auroral phenomena by recognizing that they actually fall into two groups having different patterns of occurrence (Figure 6.12). The inner one, corresponding to the luminous oval, is characterized by

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

- luminosity,
- sporadic-E on ionograms,
- spread-F on ionograms,
- soft X-rays,
- impulsive micropulsations,
- negative bays on magnetometers,
- soft but intense electron fluxes detected by satellites,
- high frequency (>4 kHz) VLF hiss, and
- rapid fading of VHF scatter signals.

In addition there is a second zone at a lower latitude, which is almost circular and covers approximately 60°–70°, with its center at about 65° geomagnetic latitude. This zone displays

- diffuse aurorae,
- radio absorption,
- sporadic-E at 80–90 km altitude,
- continuous micropulsations,
- hard X-rays of long duration,
- harder (>40 keV) electrons detected by satellites,
- VLF emissions at <2 kHz, and
- slow fading of VHF scatter signals.

This second zone of precipitation is generally thought to be connected with the outer Van Allen zone of trapped particles (Section 2.3.4); considering that it is also the outer of the two zones when plotted on a polar map, we shall call it the *outer precipitation zone*. The ionospheric phenomena in the outer zone are related to electron precipitation more energetic than that typical of the oval. The phenomena tend to be of longer duration in the outer zone. The rate of occurrence is greatest by day, whereas it is greatest by night in the oval. In both zones the phenomena are sporadic and dynamic, and both exhibit substorm behavior (Section 6.4.2). They occupy much the same latitude at midnight but become increasingly separated towards noon.

In Figure 6.6, which shows properties of electron precipitation (30 eV to 20 keV), the total flux of particles (a) is distributed like the luminous oval. However, the dayside particles being relatively soft at the higher latitudes, the total energy flux (b) has a night maximum near and just before midnight, as in the inner oval of Figure 6.12. The average energy of the particles (c) is a maximum between 60° and 70° in the morning, in the vicinity of the peak of Hartz and Brice's outer zone.

It is also interesting to compare the Hartz-and-Brice picture, made 30 years ago and based mainly on ground-based observations, with the new satellite-based results in Figure 6.13. The inner zone is the luminous intensity recorded by a

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

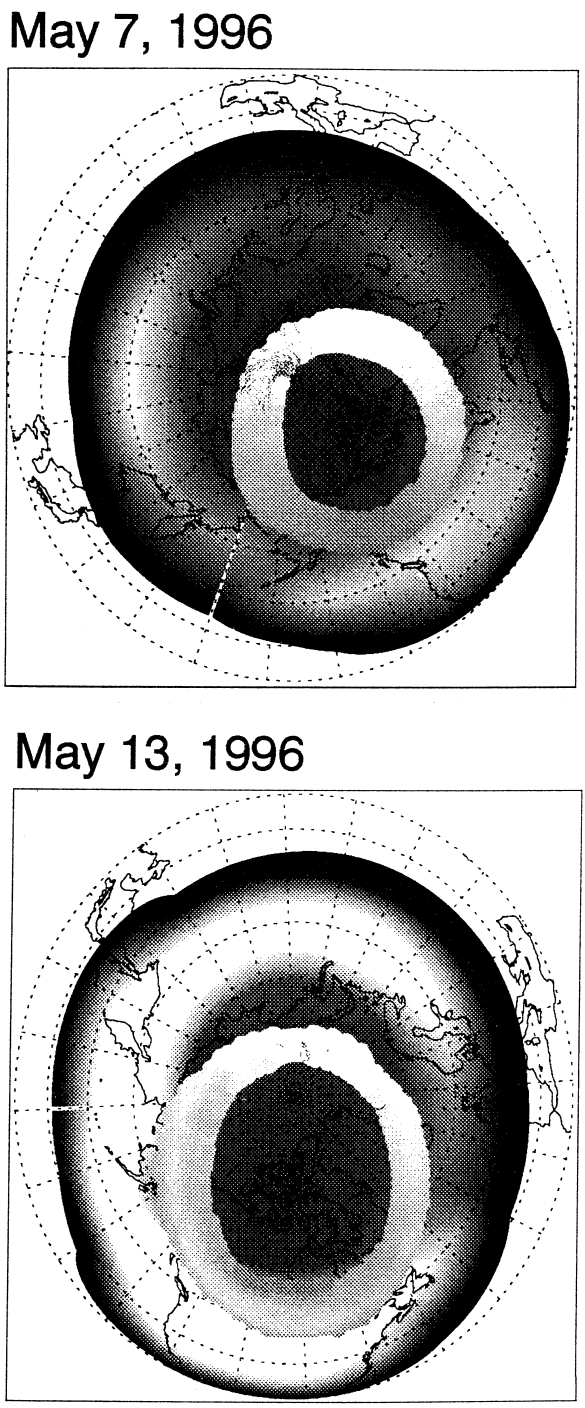


Figure 6.13. Comparison between the inner and outer precipitation zones. The auroral images were taken by the VIS camera on board POLAR, 7 May 1996 and 13 May 1996 (data courtesy of L. A. Frank, University of Iowa, USA). The radiation-belt data were obtained by the HILT electron detector on board SAMPEX (data courtesy of B. Klecker, Max-Planck-Institut für extra-terrestrische Physik, Garching bei München). Figure provided by T. I. Pulkkinen (Finnish Meteorological Institute).

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

camera on the POLAR satellite (averaged over 1 h), and the outer one is a composite of fluxes of >1 MeV electrons taken during 15 orbits of the SAMPEX satellite over the course of one day. Quiet ($A_p = 4$) and more active ($A_p = 14$) conditions are represented. On 7 May the oval is contracted and the outer zone inactive; on 13 May the oval is expanded and the outer zone intense. The plots do not confirm the morning maximum of the Hartz-and-Brice picture, but this may be because SAMPEX samples at only two local times.

Recall that the distribution of occurrence of scintillation (Section 5.3.3) also exhibits a latitudinal separation from the edge of the precipitation zone that is considerably greater by day than it is by night. This would seem to identify the region of F-region irregularity with the auroral oval rather than with the outer zone.

6.4 The substorm

6.4.1 History

As early as 1837, auroral observers had noted that during a single night there were times when the aurora was at its most intense, the activity being weaker during the periods in between (Stern, 1996). The same was true of the related magnetic signature, and it was Birkeland (1908) who first studied this tendency in magnetic records and identified what he called the "elementary polar magnetic storm." However, Birkeland's work in this area, which also involved field-aligned currents, fell into disfavor, and the topic made no further progress until the early 1960s. It was then that Akasofu and Chapman (1961), in a study of the polar disturbance (DP) field, coined the term "DP substorm" for the short periods of enhanced magnetic disturbance that Birkeland had noted more than 50 years before. Shortly thereafter, Akasofu noted that these events were often accompanied by bursts of auroral activity, which (at Chapman's insistence, it is said) were named "auroral substorms" (Akasofu, 1970). Akasofu subsequently introduced the term "magnetospheric substorm" to indicate the generality of the phenomenon and to make it clear that, although the consequences of the substorm are most apparent in the polar regions, its cause lies in the magnetosphere (Rostoker *et al.*, 1980).

6.4.2 The substorm in the aurora

The essence of a substorm as it affects the auroral regions is best described by Akasofu's analysis of the "auroral substorm" which he developed in the 1960s (Akasofu, 1968; 1977). Akasofu used all-sky camera pictures of aurorae recorded during the International Geophysical Year (1956–1958), and applied them to derive a convincing description on the global scale of the typical behavior of the luminous aurora during a substorm.

The aurora tends to be active for about an hour at a time, with quiet periods of

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

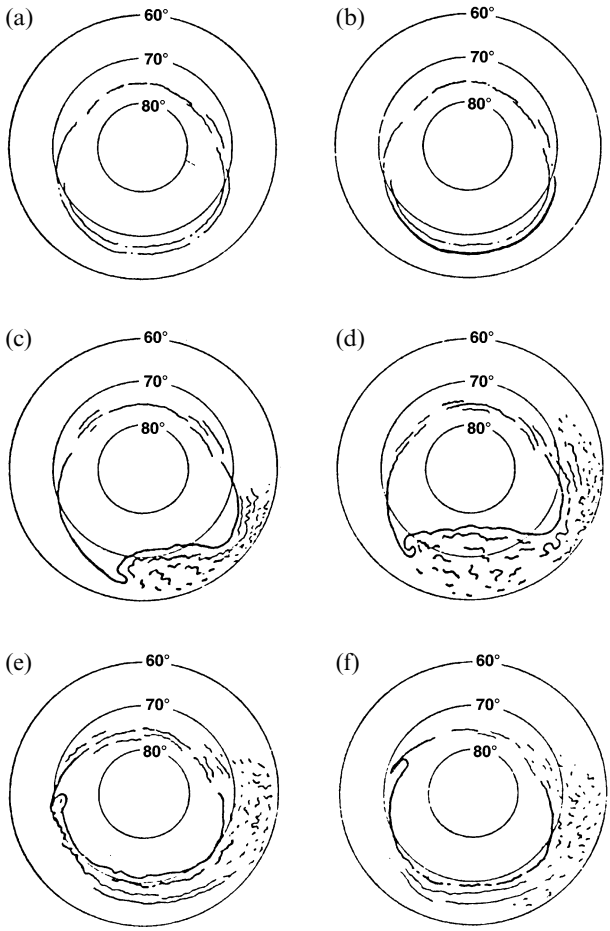


Figure 6.14. The substorm in the luminous aurora: (a) $T = 0$; (b) $T =$ 0–5 min; (c) $T = 5$–10 min; (d) $T = 10$–30 min; (e) $T = 30$ min–1 h; and (f) $T = 1$–2 h. (S.-I. Akasofu, *Polar and Magnetospheric Substorms*. Reidel, 1968, with kind permission from Kluwer Academic Publishers.)

2–3 h between, and there is also a dynamic aspect. Akasofu's representation is illustrated in Figure 6.14. The sequence begins as a quiet arc brightens and moves poleward, forming a bulge. If several arcs are present, it is often the equatorward one which brightens. Active auroral forms then appear in the bulge, equatorward of the original arc. This is called *break-up* or the *expansion phase*, and the instant when it begins is usually called the *onset*. Near midnight the oval is now broader than before, while the polar cap contained within the oval is smaller. At the same time active auroral patches move eastward towards the morning sector and other forms travel westward towards the evening. The westward movement is called the *westward-traveling surge*. After 30 min to 1 h the night sector recovers and the substorm as a whole dies away (the *recovery phase*). The sequence is likely to be repeated 2–3 h later. By defining a repeating pattern in auroral behavior it was this analysis which really established the substorm as the central concept in studies of the auroral phenomena.

The period before the break-up is now recognized as a *growth phase*, which is not so spectacular in the aurora but which was first studied in the magnetotail (Section 2.2.6), which becomes gradually more tail-like for tens of minutes to an

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

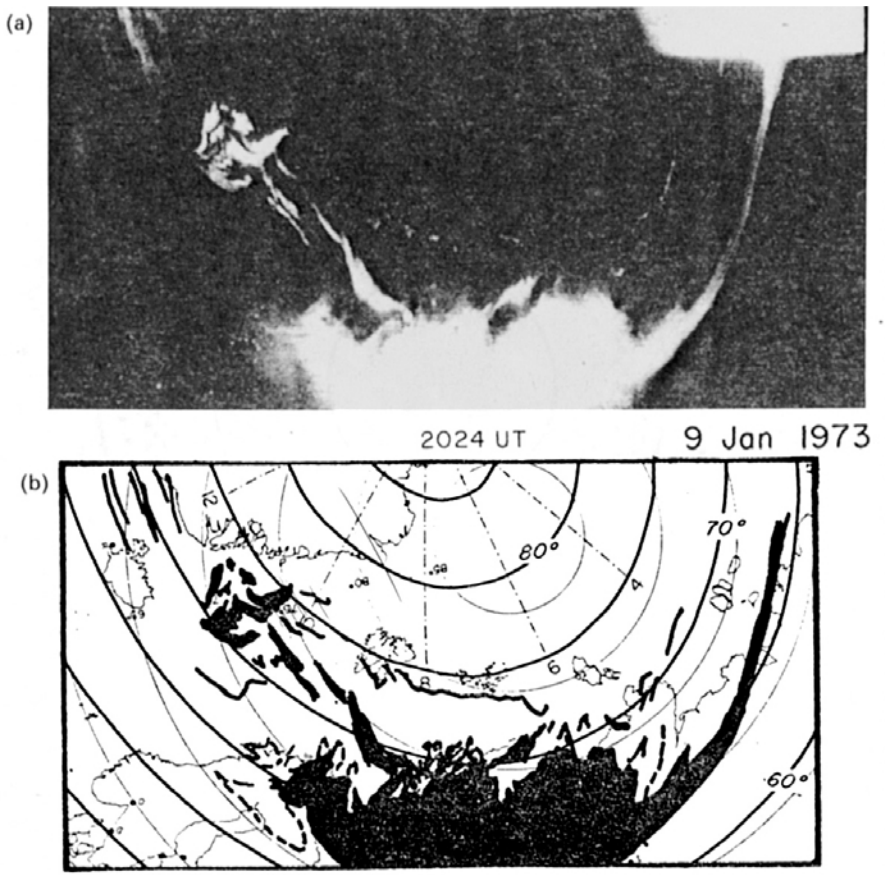


Figure 6.15. An aurora observed from space by a DMSP satellite at the maximum of a substorm, 9 January 1973 at 2024: (a) a photograph, and (b) interpretation over a map including the magnetic latitude. (S.-I. Akasofu, *Space Sci. Rev.* **16**, 617, 1974, with kind permission from Kluwer Academic Publishers.)

hour before the onset. During the growth phase the arcs of the auroral oval move equatorward and the area of the polar cap contained within the oval grows larger. The equatorward motion of the oval during the growth phase is typically several hundred m s^{-1} (Elphinstone *et al.*, 1991). Arcs form again during the recovery phase, and these also drift equatorward.

Some brightenings of the aurora do not develop into full substorms. They remain limited to a few hundred kilometers (Akasofu, 1964), and are relatively short-lived. Such events are called *pseudo-breakups*. The distinction between substorms and pseudo-breakups has been discussed by Pulkkinen (1996).

Satellite observations using downward-pointing photometers have confirmed this general picture. Figure 6.15 shows an example of a substorm breakup observed from a DMSP satellite. The auroral satellites (such as DMSP, Viking, Akebono, and POLAR) have also added much detail to the original concepts, both of the auroral oval and of the substorm – and, as so often happens, the topic

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

turns out to be more complicated than had originally been thought! For example, it now appears that the westward-traveling surge is made up of a number of localized brightenings or *surges* that do not move far as individuals. Each surge lasts for just a few minutes, and then a new surge appears to its west. Thus the aurora as a whole does indeed move westward toward the evening sector, but it goes in a series of jumps.

Murphree *et al.* (1991) have summarized the following details of the optical substorm observed by the VIKING satellite.

(1) The latitudinal width of the auroral activity does not vary systematically during the growth phase.
(2) During this phase the motion of the equatorward boundary of the diffuse aurora is generally equatorward with a speed less than a few hundred m s^{-1}.
(3) The expansion phase is preceded by auroral intensifications lasting up to several hours of local time, which fade shortly before onset.
(4) The onset region is very localized, being less than 500 km across in the ionosphere.
(5) Auroral observations under moderately active conditions indicate that auroral emissions can extend several degrees of latitude poleward of the location of the onset. This suggests that the onset region can be well away from the boundary between open and closed field-lines.
(6) When the position of the onset is mapped along the geomagnetic field to the equatorial plane, it is consistent with the location of the inner boundary of energetic particle flux, the so-called "injection boundary," observed in other studies.

In Section 6.3.5 it was pointed out that the auroral phenomena occupy not one "zone" but two. Both zones exhibit substorm behavior, and Figure 6.16 shows an overall picture illustrating how the substorm develops in each zone, as represented by the fluxes of softer (5 keV) and harder (50 keV) electrons, respectively. The oval and the outer zone are obviously related to each other in some way. The most likely mechanism is that, when the auroral oval is active in a substorm, the outer zone becomes populated with energetic particles that drift in longitude and are subsequently precipitated. However, not all the physical connections between the two zones have been explained fully.

6.4.3 Ionospheric aspects of the substorm

The enhanced precipitation of energetic electrons during a substorm increases the rate of ionization of the ionosphere, and of the lower ionosphere in particular, by an amount depending on the particle flux, and over a range of altitudes determined by the particle energies (Figure 2.26). Consequently, the substorm behav-

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

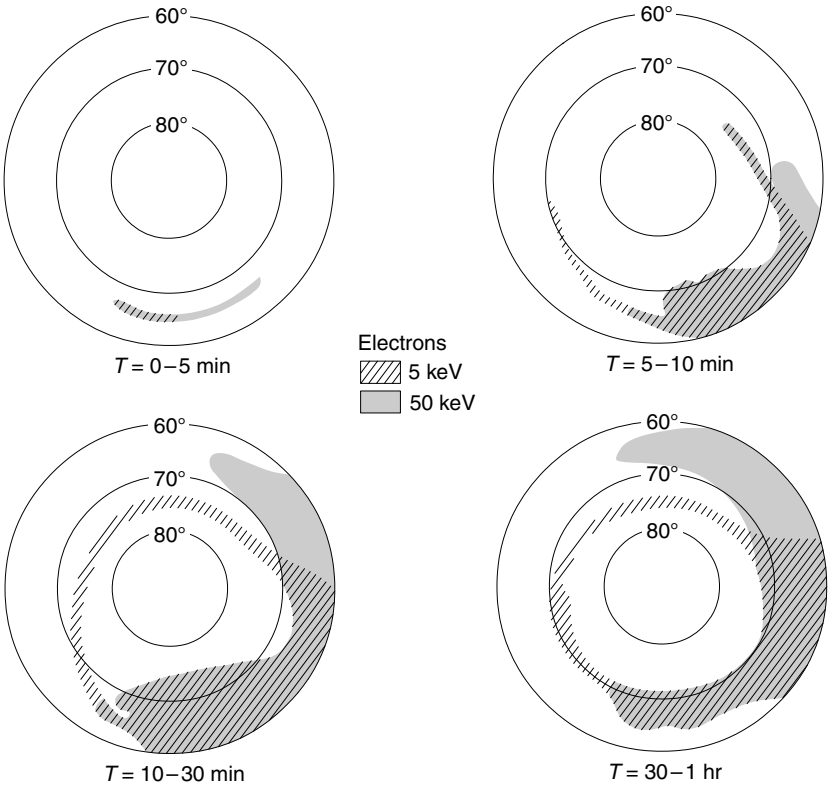


Figure 6.16. The typical development of electron precipitation in a substorm. Note that the two zones are distinct on the day side. (S.-I. Akasofu, *Polar and Magnetospheric Substorms*. Reidel, 1968, with kind permission from Kluwer Academic Publishers.)

ior observed in the luminous aurora carries over into the various ionospheric effects (Section 6.3.4). The E-region reflections called radar aurora, radiowave absorption in the D region, X-ray generation, and the occurrence of magnetic bays are all, therefore, substorm phenomena. Some of these will be described in more detail in Section 6.5 and Chapter 7.

6.4.4 Substorm currents

Figure 6.17 shows one of the earlier descriptions of the current flowing during an individual substorm. This is still an *equivalent-current system*, because it assumes that the current flows only horizontally. Note that the intensity is relatively greater on the morning side of midnight. It will be seen that Figure 6.17 is considerably different from Figure 2.23, which showed the auroral electrojets converging on the Harang discontinuity at midnight. The patterns of Figures 6.17 and 2.23 are related rather as is the auroral oval to the auroral zone.

However, the acceptance of field-aligned, or Birkeland, currents (Section 2.3.6) has fundamentally changed the approach to current modeling, because currents

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

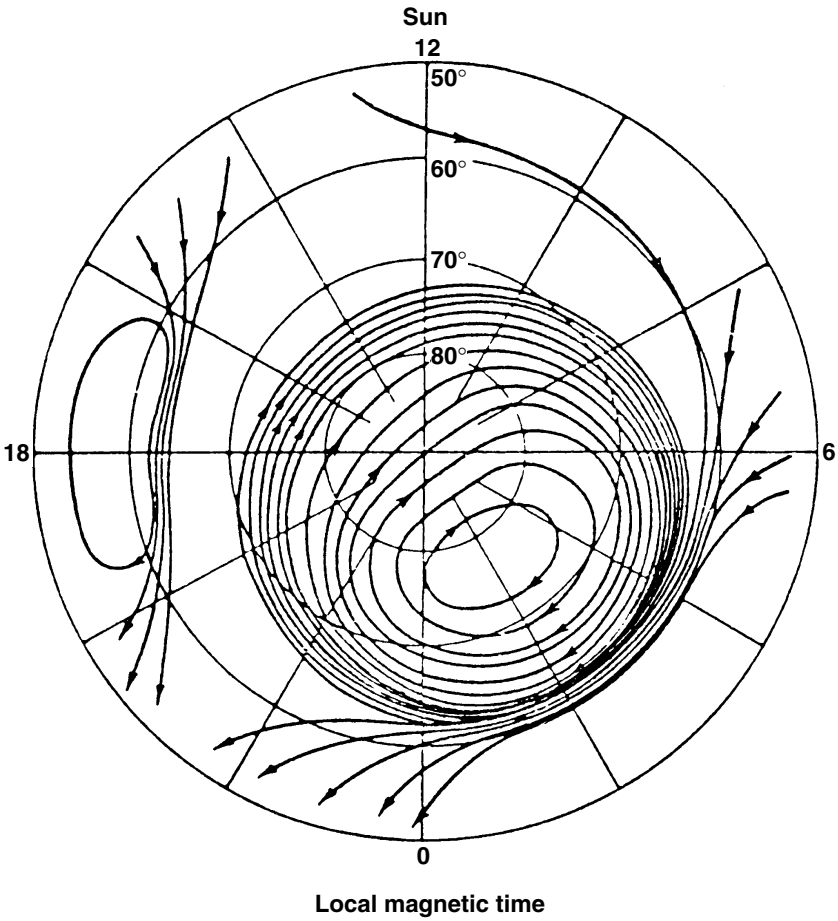


Figure 6.17. The equivalent current system of a magnetic substorm. The concentrations of current lines in the early morning and near 1800 LT would appear as electrojets. (S.-I. Akasofu and S. Chapman, *Solar–Terrestrial Physics*. By permission of Oxford University Press, 1972.)

within the magnetosphere as well as currents flowing between the magnetosphere and the ionosphere may be included in the circuit. Despite general agreement on this point, the form of the current system during a substorm is still a topic of investigation.

One influential concept in present-day modeling of substorm currents is that of the *current wedge*. As we shall see, the magnetotail collapses in a limited region when a substorm begins, and the cross-tail current from that region becomes diverted along field-lines (as Birkeland currents) into the ionosphere. There the circuit is completed in the E region, probably by an electrojet flowing along an arc or through some other form whose conductivity is enhanced by particle precipitation. Figure 6.18(a) shows the magnetospheric part of the circuit, and Figure 6.18(b) the "substorm electrojet" in the ionosphere. The electrojet, which flows

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

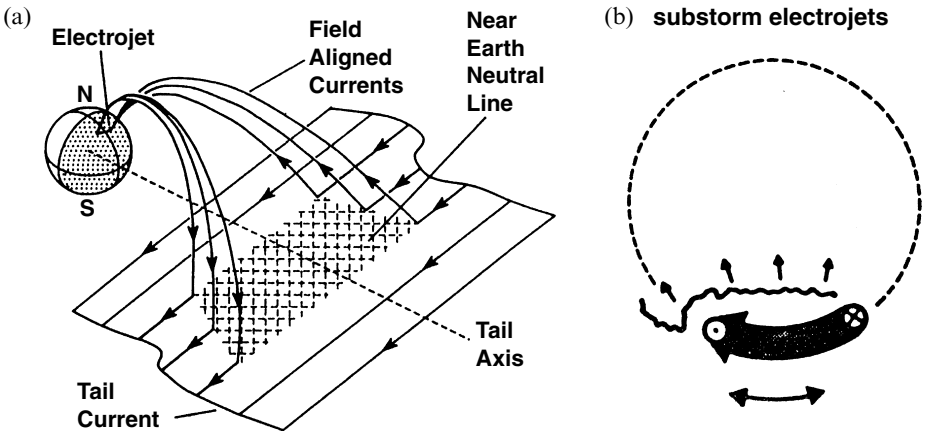


Figure 6.18. (a) The substorm current wedge due to the diversion of tail current to the ionosphere. (Y. Kamide, Report ESA SP-389, 1996, after McPherron *et al.*, 1973.) (b) The substorm electrojet in the auroral zone. (G. Rostoker, in *Magnetospheric Substorms*, copyright by the American Geophysical Union, 1991.)

westward in the midnight sector, connects to an upward field-aligned current at its western end, which also concides with a bright auroral feature.

However, this is only part of the picture. The local collapse in the magnetotail at the onset of a substorm accelerates particles towards the Earth, and some become trapped to form a partial ring current (Section 2.3.5) that is completed by Birkeland currents to the ionosphere and currents within the ionosphere. These are driven, at least in part, by the electric field due to the general polar convection (Sections 2.4.1 and 2.4.3), which is likely to be enhanced during substorm activity. Various suggestions about the form and relationship of these currents flowing during a substorm have been put forward. Figure 6.19 indicates one possibility; Figure 6.19(b) shows the "convection electrojet" in the ionosphere.

Kamide (1996) has pointed out that, whereas the enhanced conductivity of the ionosphere is the main factor controling the westward electrojet near midnight, the eastward electrojet which flows in the late evening sector, before the Harang discontinuity, is dominated by a northward electric field (in the northern hemisphere). An electric field, this time southward, also dominates the situation in the westward electrojet later in the morning (Figure 6.20). These are the fields generated by plasma convection. The components due to the wedge of current and convection do not vary in the same way during a substorm, however – their typical time constants are 15 min and 2 h, respectively (Rostoker, 1991) – so that the total behavior cannot be expected to be simple, or even the same in all cases. The total electrojet should be a combination of Figures 18(b) and 19(b), but in varying amounts.

Although the electrojet is usually conceived in terms of a single wedge of current, more recent studies (Rostoker, 1991) have shown that it is composed of a sequence of short bursts (lasting about 12 min) of westward current (christened *wedgelets*!), following one after the other and often appearing sequentially further to the west. Thus there is a gradual westward progression of the current, as there is of the luminosity (Section 6.4.2).

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

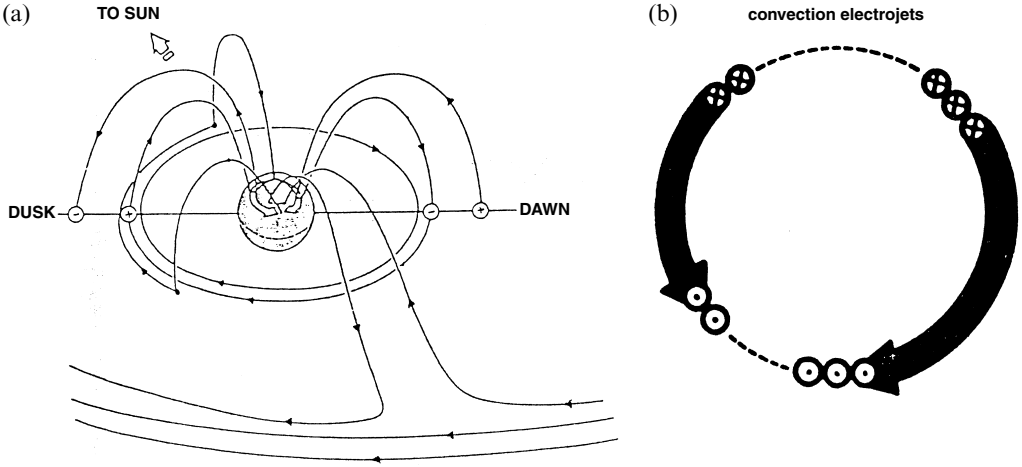


Figure 6.19. (a) Magnetospheric currents showing the ring current and associated Birkeland currents. (Y. I. Feldstein, in *Magnetospheric Substorms*, copyright by the American Geophysical Union, 1991.) (b) The substorm electrojet in the auroral zone. (G. Rostoker, in *Magnetospheric Substorms*, copyright by the American Geophysical Union, 1991.)

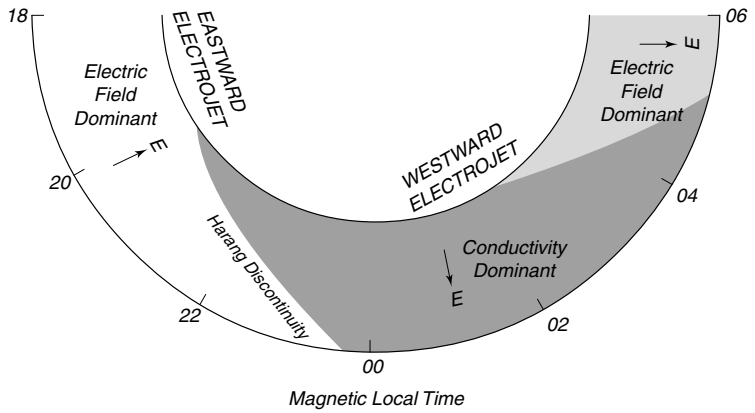


Figure 6.20. Regions of the electrojet dominated by conductivity and by electric fields. (Y. Kamide, in *Auroral Physics*, Cambridge University Press, 1991, p. 385.)

6.4.5 The substorm in the magnetosphere

If the substorm is the unit of auroral activity, then it is important to discover the details of the phenomena through which the substorm is revealed, and this includes its appearence in the magnetosphere. Moreover, if we are to predict when the high-latitude ionosphere is likely to be affected by substorms we have to understand the essential nature of the substorm and the factors that make it happen; these factors concern the dynamics of the magnetosphere and its interaction with the solar wind. It is another topic that is still being actively researched, both experimentally and theoretically. Although the final story has not yet emerged, there are some aspects that seem well established.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

Field-line circulation

The circulation of the magnetosphere, discussed in Section 2.4, is mainly driven by magnetic merging on the sunward side of the magnetopause, but its continuity depends on reconnection in the plasmasheet which lies along the central plane of the tail. If we select a field-line at high latitude on the day side of the Earth and follow its progress, we find that it goes through a sequence:

(1) connecting with the IMF which divides it in two;
(2) convecting over the north and south poles as two separate halves;
(3) reconnecting in the tail; and
(4) returning to a more dipolar form and returning to the day side.

In a steady state these stages would be in balance. Substorms occur because neither the dayside connection nor the nightside merging are continuous processes. Thus, energy accumulates in the tail and the substorm marks its sudden release.

The Interplanetary Magnetic Field (IMF) reaching the Earth lies principally in the ecliptic plane, but it generally has some northward or southward component as well, and it couples most strongly with the geomagnetic field when that component is southward (Section 2.4.2). When the IMF turns from northward to southward, the connection rate increases. For a while more open field is produced than removed, the total magnetic flux in the polar cap increases, the auroral oval moves equatorward, and the tail of the magnetosphere grows fatter, representing a store of energy. This energy is released in the substorm, when the reconnection rate in the tail exceeds the supply of magnetic flux coming from the polar regions, the tail becomes more dipolar, flux is lost from the polar caps and the auroral oval shrinks again. This sequence of events in the magnetosphere may be identified with the phases of the auroral substorm observed from the ground: the growth phase, the expansion phase, and the recovery phase.

Behavior in the tail

In the magnetosphere the growth phase corresponds to an increase in erosion from the front of the magnetosphere, and the plasma sheet and the current sheet become thinner (though not necessarily at the same time) as illustrated in Figure 6.21(a).

One concept of the expansion phase begins with the formation of a neutral line nearer to the Earth than that which exists during quiet times (Figure 6.21(b)). Here the magnetotail collapses because the magnetic field has gone to zero in a localized region, and the cross-tail electric current is diverted into the current wedge as described above. In fact the collapse and the diversion of current must go together because of the Biot–Savart law. At the same time, satellites at geosynchronous distance observe an increase in the flux of energetic electrons, and the geomagnetic field becomes more dipolar.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

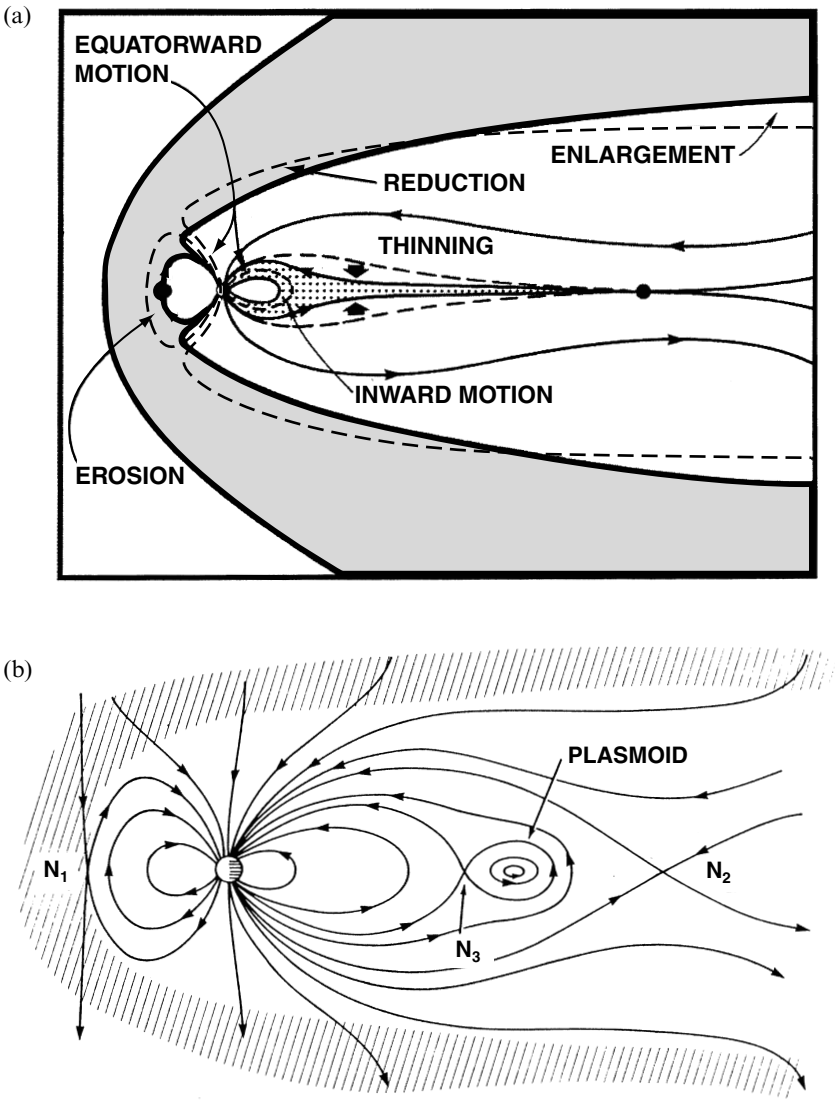


Figure 6.21. (a) Changes in the magnetosphere during the growth phase of a substorm. (R. L. McPherron *et al.*, *J. Geophys. Res.* **78**, 3131, 1973, copyright by the American Geophysical Union.) (b) The near-Earth neutral line (N_3) and the plasmoid formed in a substorm. N_1 is the merging point at the front of the magnetosphere and N_2 the merging region in the distant tail. (D. P. Stern, *Rev. Geophys.* **34**, 1, 1996, copyright by the American Geophysical Union.)

The region of the tail between the two neutral lines forms a *plasmoid*, which is ejected along the magnetotail as the recovery phase begins; this may be detected by satellites (20–100)R_E down the tail as a burst of energetic particles moving away from the Earth. Satellites near the neutral sheet detect a loss of particle flux during the expansion phase, indicating that the plasma sheet becomes thinner at that time.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

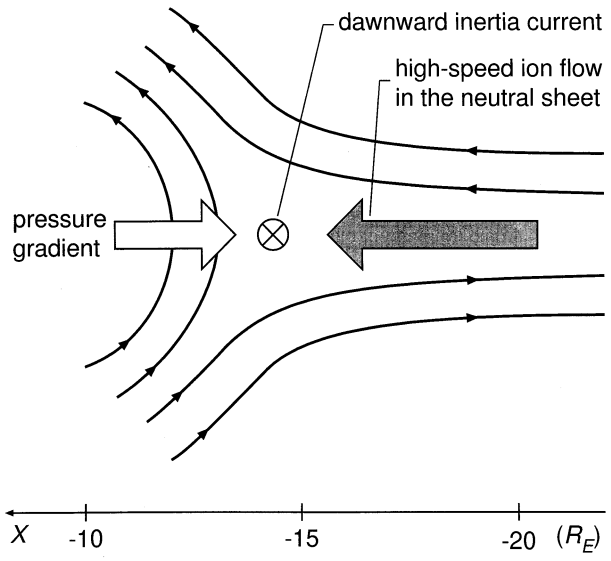


Figure 6.22. A Y-type neutral point at the Earthward edge of the neutral sheet where the high-speed ion flow is stopped by the dipole-like field of the inner magnetosphere. (K. Shiokawa *et al.*, *Geophys. Res. Lett.* **24**, 1179, 1997, copyright by the American Geophysical Union.)

Quite recently, spacecraft in the magnetotail (AMPTE and GEOTAIL) have observed *bursty bulk flows* (*BBFs*) in the plasma sheet. Plasma in the sheet generally flows at speeds of less than 100 km s^{-1}, but during a BBF, which typically lasts for about 10 min, the speed exceeds 400 km s^{-1} and the direction of flow is Earthward. BBFs occur at all observed distances beyond $15R_E$, and appear to be associated with the occurrence of substorms (Angelopoulos, 1996). If they are observed a long way down the tail (e.g. at $(90–100)R_E$) the event appears some 90 min after the substorm, suggesting that there is a centre of acceleration that retreats progressively down the tail away from the Earth. When the fast-flowing plasma gets closer to the Earth, it is stopped by the stronger geomagnetic field which is dipolar in form. Figure 6.22 illustrates the stopping region. Note that the magnetic field is configured with a Y-type neutral line (compare with Figure 2.20). In this treatment the flow stops at the boundary between the tail field and the dipole field, and this region is also the inner edge of the plasma sheet.

Various theories

The exact configuration of the magnetotail during the phases of a substorm has not been established fully, and several models – among which the observations have not yet been able to distinguish – have been put forward. Some models involve a local reversal of the tail field, and others include multiple neutral lines to correspond to the multiple arcs seen in aurorae. Liu (1992) has summarized the contending theories as six models.

(a) Formation of a neutral line in the magnetotail at a distance of $(10–20)R_E$, which allows magnetic reconnection between the lobes of the magnetotail.

(b) Generation of the Kelvin–Helmholtz instability (Section 2.5.6) in the mag-

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

netospheric boundary layer, by enhanced reconnection at a neutral line some $100R_E$ distant down the tail.

(c) A "thermal catastrophe" in the plasma sheet, due to the sheet becoming opaque to Alfvén waves and consequent sudden heating.

(d) Intense field-aligned currents due to an increase in the rate of field reconnection on the day side of the magnetosphere, leading to the "current wedge" of the substorm and a collapse of field in the magnetotail.

(e) A disruption of the cross-tail current due to a current instabilty.

(f) A "ballooning instability" invoking a transition between field configurations that are essentially dipolar and essentially tail-like, which again diverts the cross-tail current.

These are illustrated in Figure 6.23, but nothing would be gained by going into all their details here. The task of making a synthesis of all the various substorm observations and theories has been considered by Elphinstone *et al.* (1996).

6.4.6 The influence of the IMF and the question of substorm triggering

The magnetic power of the solar wind

It is clear that the term "substorm" includes a considerable range of phenomena, but the central idea is of a sudden and sporadic episode in the magnetosphere in which a large amount of stored energy is released. The energy comes initially from the solar wind, and important factors in the occurrence of substorms, therefore, are the energy flux of the solar wind and the efficiency with which the energy is coupled into the magnetosphere. It is found that the index AE, which indicates the level of geomagnetic activity in the northern auroral zone, is well correlated to a quantity

$$\varepsilon = vB^2 \sin^4(\theta/2)\, l_0^2, \tag{6.2}$$

where v is the solar wind speed, B is the magnitude of the IMF, l_0 is a length related to the cross-section of the magnetosphere ($7R_E$), and θ is the "clock angle" of the IMF seen from the Earth (as defined in Section 5.1.2.).

The magnetic energy reaching the magnetosphere per unit time is proportional to $vB^2 l_0^2$: this is the "magnetic power" of the solar wind. The expression $\sin^4(\theta/2)$ is intended to represent the fraction of this power coupled into the magnetosphere. Its form gives a gradual transition between full coupling when the IMF is fully southward ($\sin^4(\theta/2) = 1$) and zero coupling when it is fully northward ($\sin^4(\theta/2) = 0$). If $B_z \ll B_y$, $\theta/2 = 45°$, and the coupling factor is 0.25. Some other expressions based on different combinations of solar-wind parameters also correlate to the occurrence of substorms, though ϵ is perhaps the best (Figure 6.24).

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

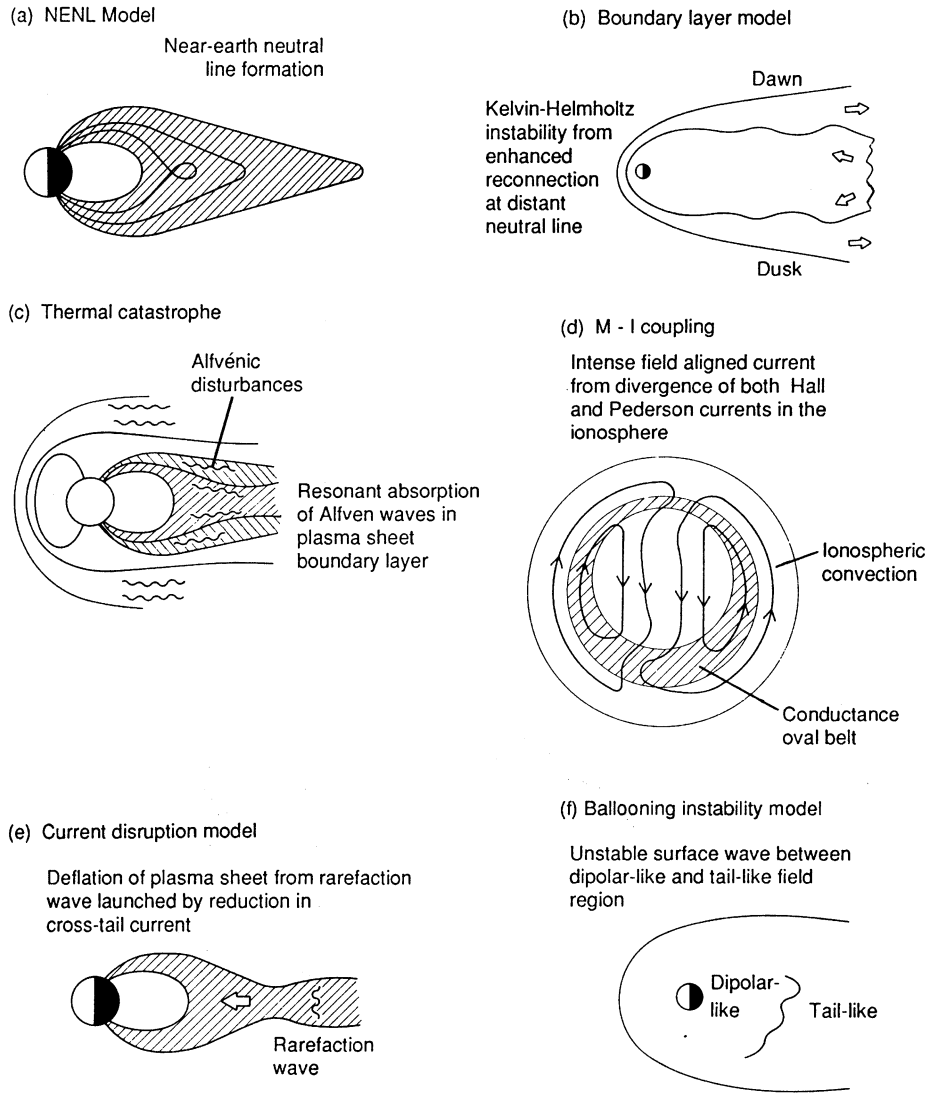


Figure 6.23. A selection of substorm-triggering ideas. (A. T. Y. Lui, in *Magnetospheric Substorms*, copyright by the American Geophysical Union, 1991.)

The influence of B_z on triggering

Although much is known about the sequence of events in a substorm, it is not clear what causes the event to begin. Clearly, a store of energy must have built up awaiting release, but then we still have to ask whether the substorm is triggered by some other identifiable event, for instance in the solar wind, or whether it might be a spontaneous phenomenon without apparent cause. This point is clearly a vital element in any substorm theory, and (at the time of writing) it is not yet settled.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

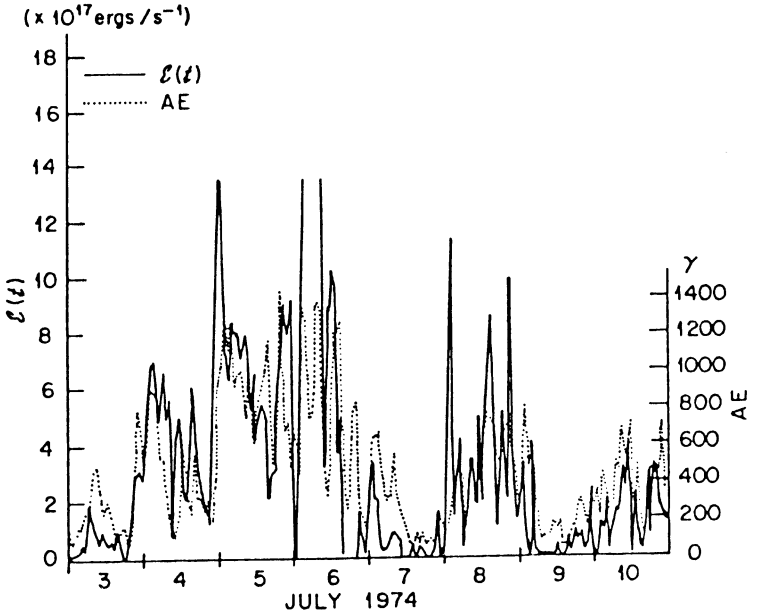


Figure 6.24. Correlation between the AE index of magnetic activity and the parameter ε during a storm in July 1974. (Reprinted from S.-I. Akasofu, *Planet. Space Sci.* **27**, 425, copyright 1979, with permission from Elsevier Science.)

Some things can be said, however. That substorms occur most frequently when B_z is southward has been known for many years, and it is found that the beginning of the substorm often coincides with a southward turning of the IMF. However, there are also cases when the substorm begins as the IMF turns northward, having previously been southward for an hour or two. In such a case it appears that a southward IMF puts energy into the magnetosphere and then the shock of the northward turning in some way triggers its release. Indeed, in many cases the growth phase begins as the IMF turns south. It is possible that there may be more than one kind of trigger.

The substorm rate

If the speed of the solar wind is 440 km s^{-1} (Figure 2.2(a)), the substorm rate is in the range 800–1500 per annum, or about 2–4 a day on the average (Borovsky *et al.*, 1993; N. Flowers, private communication). Also, the rate of occurrence of substorms increases with the speed of the solar wind. Observations of substorms in auroral absorption give a median rate of 3.8 per day at 450 km s^{-1} increasing to 7.2 per day at 700 km s^{-1}, approximately a v_{sw}^2 dependence (Hargreaves, 1996).

6.4.7 Relations between the storm and the substorm

Storms and substorms are defined in different ways, the former mainly from magnetic observations at low latitude where the greatest influence is the ring current,

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

and the latter from observations at a higher latitude where the greatest contribution comes from the auroral electrojet. Their effects at the ground are usually represented by the magnetic indices D_{st} and AE (Sections 2.5.2 and 2.5.4). It is well known that during a significant storm there will almost certainly be one or more substorms. On the other hand, substorms may very well occur when there is no storm.

Because it is the more frequent occurrence, the substorm has usually been regarded as the fundamental element, which, it is commonly supposed, leads to an increase in the population of trapped particles in which an increased ring current may then flow. This view is supported by direct observations in the magnetotail, in which a significant difference between those substorms occurring during a storm and those occurring at other times is found (Baumjohann, 1996). In a "storm-time substorm" the magnetic field moves from tail-like to dipole form in a matter of 15–30 min, whereas in "non-storm substorms" the change is both slower and less complete. There is a greater reduction of magnetic pressure in the storm-time substorm and the ion temperature in the tail is larger throughout. These results suggest that there are two kinds of substorm (perhaps arising at different distances down the tail), one of which is the more effective at populating the ring current (and thus promoting the signature of the classical magnetic storm).

There are also contrary arguments, notably from studies showing that the ring current is as likely to grow before the auroral activity begins as it is to follow it. Furthermore, there is a good correlation between the magnitude of the ring current and the electric field across the magnetosphere, showing that the solar wind affects the magnetic storm directly, as well as indirectly via substorm activity (Clauer and McPherron, 1980).

Plainly, the nature of the storm–substorm relationship is not yet fully understood.

6.5 The E region at high latitude

6.5.1 Introduction

At middle latitude the E region is easily the most boring part of the ionosphere. It behaves as an α-Chapman layer (Section 1.3.3) and supports the (S_q) current generated by atmospheric tides. Sporadic-E (Section 1.4.2) adds some interest, but beyond that there is little more to be said. The same is true at high latitude while geophysical conditions are quiet, but, when the Sun is active, the high-latitude E region becomes arguably one of the most exciting parts of the ionosphere. It then differs markedly from the mid-latitude and equatorial E regions – in terms of ionization sources, plasma processes, and radio-propagation characteristics. It is often the case that precipitating particles are the dominant source of ionization.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

When thus enhanced the E region supports the auroral electrojet, in which instabilities may arise. Ionograms exhibit sporadic-E of the auroral variety.

6.5.2 The polar E layer

The most benign part of the high-latitude E layer is over the polar cap – that is, poleward of the auroral zone. Here it is essentially under solar control; it varies with the solar zenith angle and exhibits strong seasonal effects, as does the mid-latitude E region.

6.5.3 The auroral E layer under quiet conditions

When K_p is small the auroral oval retreats poleward (see Figure 6.7), and, under quiet conditions, there is relatively little disturbance by precipitating particles in the nominal auroral zone, from, say, 60° to 70° magnetic latitude. The normal ionospheric layers occur much as at middle latitude and are subject to the same diurnal, seasonal, and sunspot-cycle variations. Figure 6.25 shows typical electron-density profiles from the Chatanika incoherent-scatter radar (ISR) compared with ionograms from College, Alaska, sites both near 65° magnetic latitude.

These results are for magnetically quiet conditions near sunspot minimum. Because of its high latitude the site was illuminated by the Sun throughout the day. Thus, even the ionogram at 0215 LT shows a strong E layer that masks the F region, giving the G condition (Section 5.2.1). Observing by ionosonde alone would suggest, falsely, that the F layer was absent. Note that $T_e > T_i$ in the F region, which is usual.

6.5.4 The disturbed auroral E layer

The main disturbances affecting the auroral E layer are geomagnetic storms and substorms (Section 6.4). With increasing activity, and particularly if $K_p > 3$, the auroral oval expands both poleward and equatorward, and the auroral structure and motion become much more dynamic.

The precipitating electrons of energy 1–10 keV which create the visual aurora also create the auroral-E ionization. As pointed out in Section 2.6, fast electrons (and protons) entering the atmosphere produce one ion pair (an ion plus an electron) for each 36 eV of energy lost, most of which is deposited towards the end of the path. Since the average ionization potential of the ionospheric atoms and molecules is about 15 eV, approximately 40% of the energy goes into ionization and 60% goes into the motion of the electron produced, which subsequently thermalizes. In the E region the neutral air is dense relative to the higher levels, and the recombination between electrons and ions proceeds rapidly.

Altitude profiles of the rate of ionization due to a flux of 10^8 electrons $\mathrm{cm}^{-2}\,\mathrm{s}^{-1}$ at several initial values of energy E_p (keV) precipitating along geomagnetic field

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

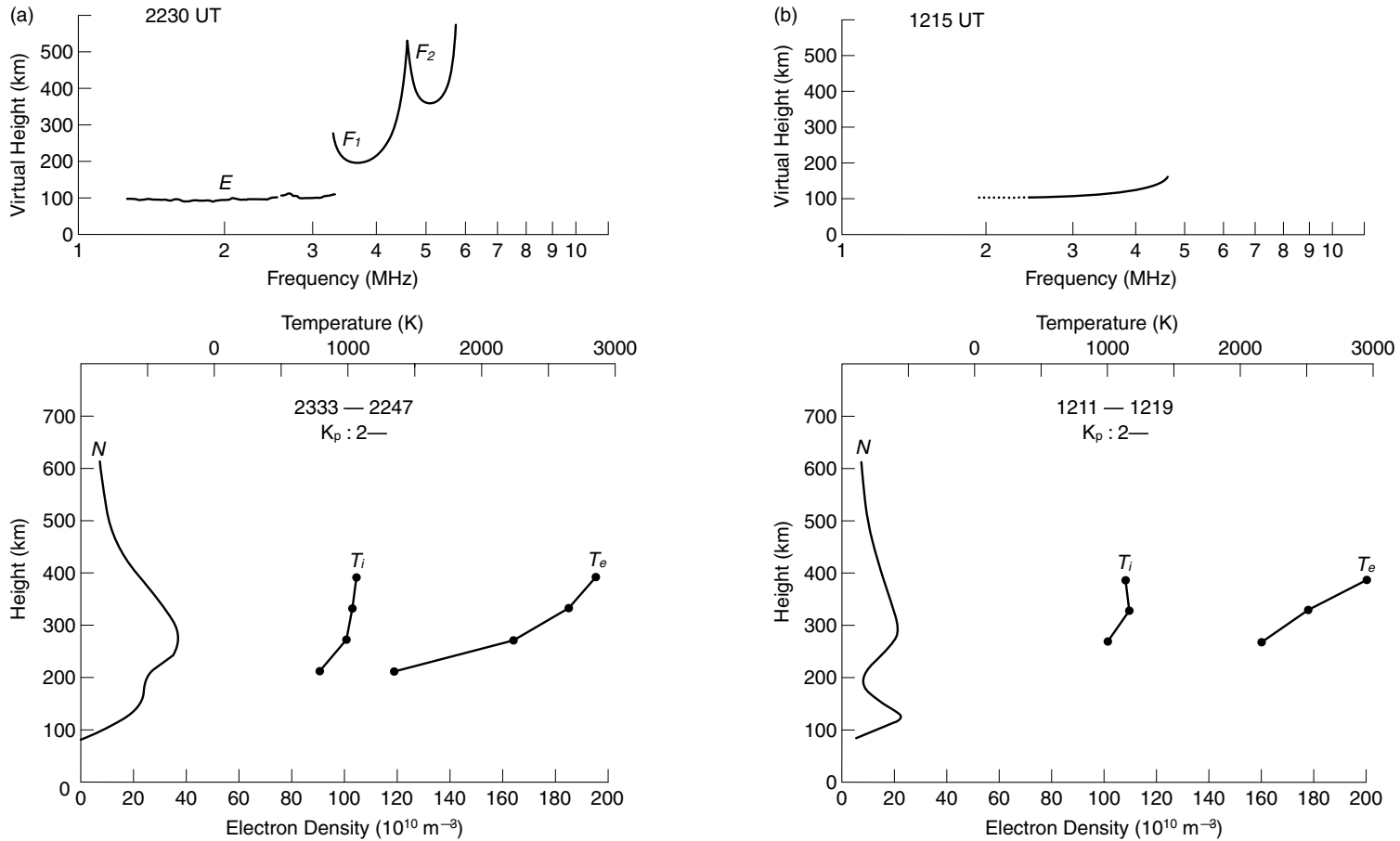


Figure 6.25. Comparable electron-density profiles from the Chatanika incoherent-scatter radar and ionograms from the College ionosonde, both in Alaska, 16 July 1971 Ion and electron temperatures are also given: (a) afternoon (Alaskan time being UT – 10 h), and (b) night. Note that, by night, the electron density is reduced, the valley between the E and F regions is more marked, and region F is masked by E, giving a simpler ionogram. (H. F. Bates and R. D. Hunsucker, *Radio Sci.* **9**, 455, 1974, copyright by the American Geophysical Union.)

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

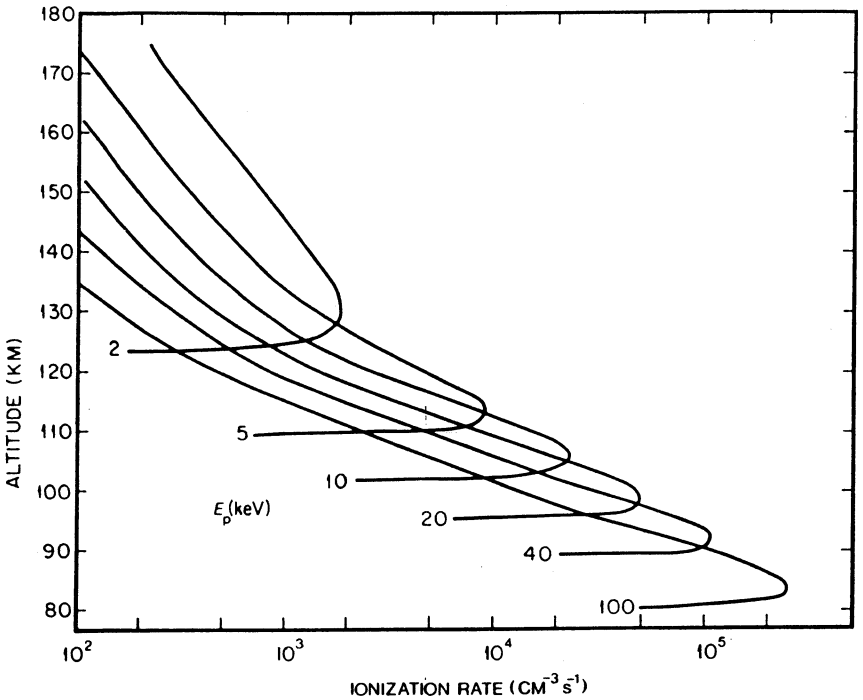


Figure 6.26. Profiles of the rate of ionization due to monoenergetic fluxes of 10^8 electrons cm^{-2} s^{-1}, incident from above, for the energy range (2–100 keV) having greatest effect in the E (and D) regions. (M. H. Rees, *Physics and Chemistry of the Upper Atmosphere*, Cambridge University Press, 1989.)

lines into the auroral ionosphere are shown in Figure 6.26. It is instructive to compare these ionization profiles with luminosity profiles (for example Figure 6.8), as a general indication of the energies of the particles which cause the auroral luminosity and the disturbed E region.

The regions of enhanced electron density are also regions of high conductivity and this is where the auroral electrojet flows, its magnitude increasing with the auroral luminosity. The form of these current systems is discussed in Sections 2.5.3 and 6.4.4. Plasma waves generated in the electrojet produce the various types of radar-backscatter signature discussed in Section 6.5.5.

The relation between the visual aurora and auroral-E ionization has been studied in a semi-quantitative fashion by Hunsucker (1975). Figure 6.27 shows simultaneous all-sky camera, ionosonde, and incoherent-scatter radar data during the passage of an auroral arc through the fields of view of the instruments. The E-region electron density is greatly increased within the arc.

The rapid changes of electron density that may be observed by incoherent-scatter-radar over an hour are illustrated in Figure 6.28. Note that, during the electron-density spike, the ionogram shows intense sporadic-E. The enhancement of the E region may be surprisingly large during major disturbances. During the

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

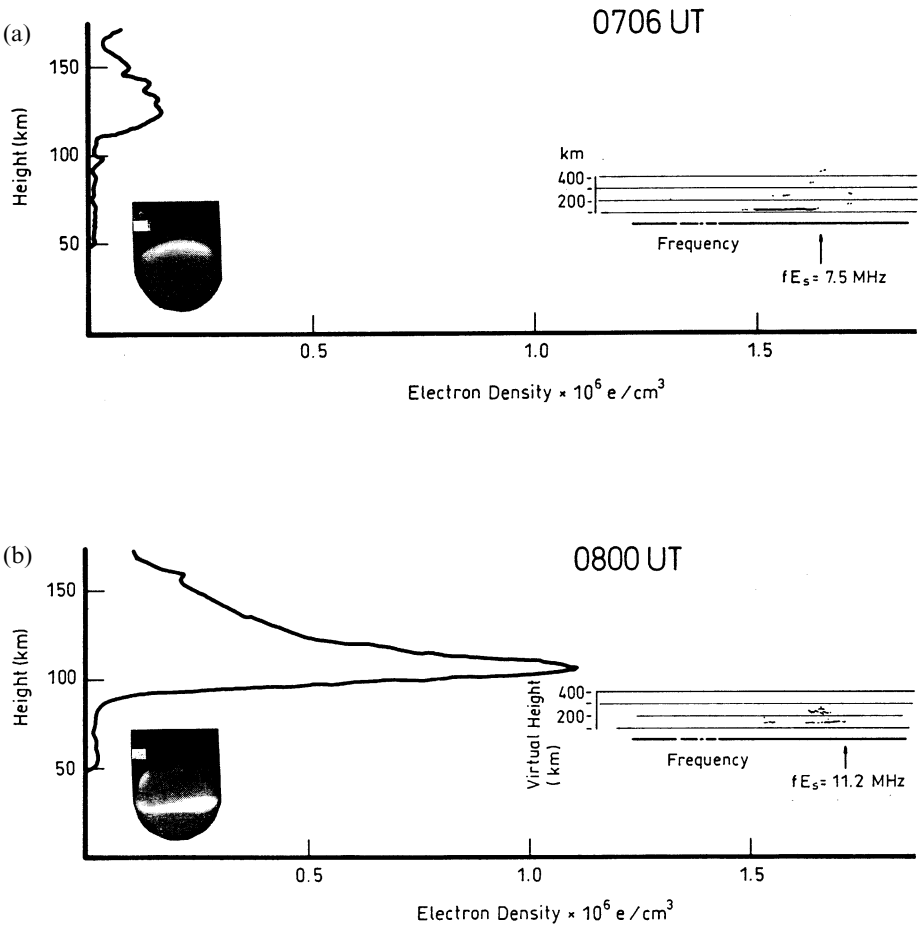


Figure 6.27. E-region electron density profiles and ionograms associated with the passage of an auroral arc across the field of view, 2 March 1973. The incoherent-scatter radar was at Chatanika, Alaska, and the all-sky camera and ionosonde were at College nearby. The radar beam was just south of the arc in (a) but in the centre of the arc in (b). The maximum electron density and the penetration frequency of the E layer were greatly increased within the arc. It was late evening. (R. D. Hunsucker, *Radio Sci.* **10**, 277, 1975, copyright by the American Geophysical Union.)

great magnetic storm of August 1972 the E-region electron density exceeded 1.8×10^6 cm^{-3} (Figure 6.29), one of the higher values of electron density observed in the E region.

The morphology, structure, and dynamics of the auroral-E layer have been described in some detail by Hunsucker (1975) and others referred to therein.

6.5.5 Auroral radar

Most of our knowledge of irregularities in the auroral E layer is based on data obtained by direct backscatter from field-aligned irregularities within the auroral

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

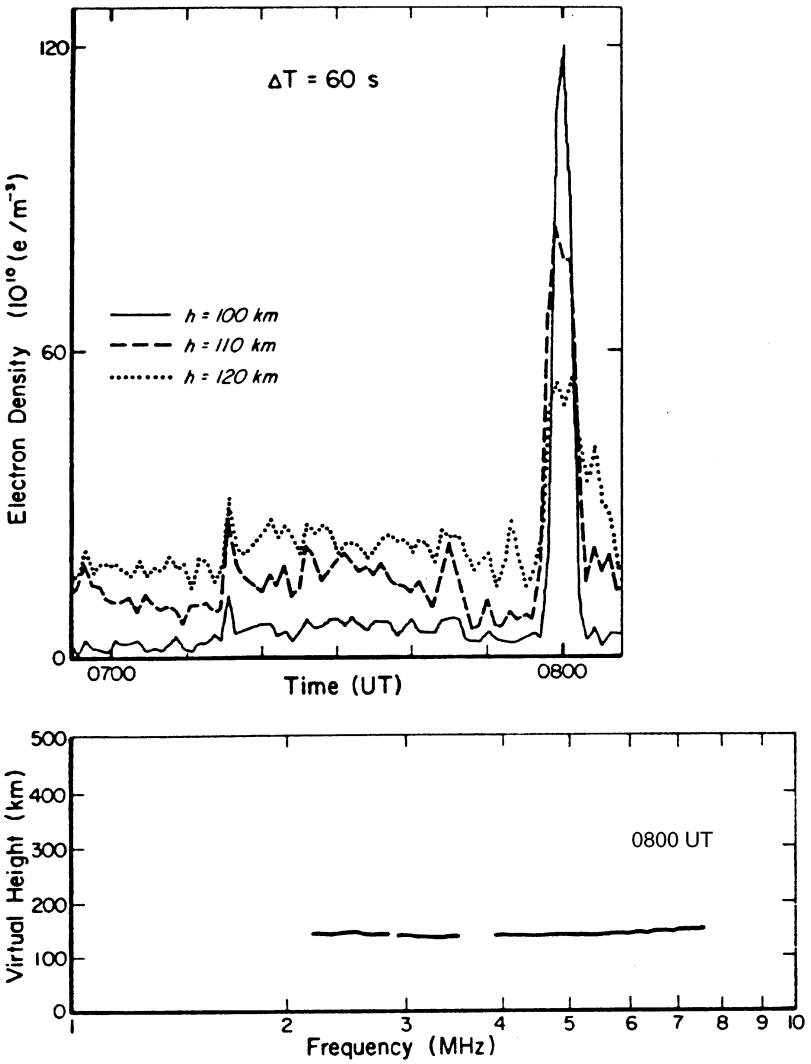


Figure 6.28. The variation of electron density at three heights, showing a sharp spike in the electron density at 0800 UT (2200 LT) on 2 March 1973. At the same time the ionosonde registered sporadic-E above 7 MHz. Observations from Alaska, using the Chatanika ISR and the College ionosonde. (H. F. Bates and R. D. Hunsucker, *Radio Sci.* **9**, 455, 1974, copyright by the American Geophysical Union.)

oval using radars operating in the VHF/UHF frequency range. The term *radio aurora* is often used for the phenomena so observed, and *radar aurora* is equally valid. There is now a voluminous literature on the subject.

There is an enormous difference between the scattering cross-sections for coherent and incoherent radar, the former being the stronger by 50–80 dB. Backscatter from auroral E-layer irregularities has been classified into four groups in terms of their line-of-sight Doppler velocities, as shown in Figure 6.30, which also summarizes their essential properties.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

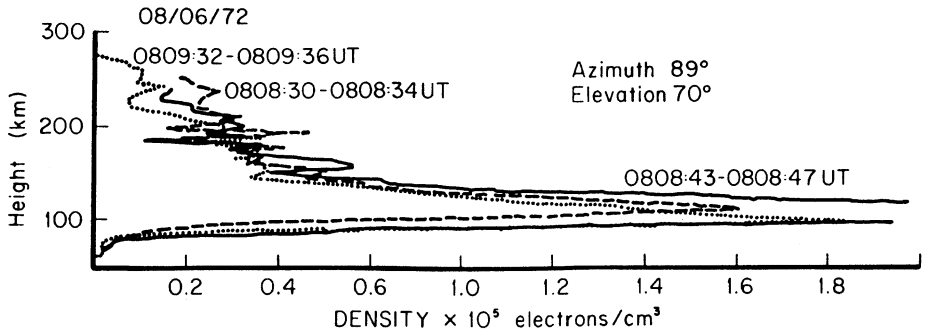


Figure 6.29. An electron-density profile of the auroral E region during a great magnetic storm in August 1972. (R. D. Hunsucker, *Radio Sci.* **10**, 277, 1975, copyright by the American Geophysical Union.)

Theory

The two most generally accepted theories to explain the observations are the *two-stream plasma instability* and the *gradient-drift mechanism*. These plasma instabilities are generated in the auroral electrojet and produce electrostatic ion waves that may scatter incident radio waves as discussed in Section 3.5.1. A necessary condition for the occurrence of these instabilities is a sufficiently large relative velocity between the electrons drifting in an electric field and the ions whose motion is dominated by collisions. The waves travel nearly perpendicular to the geomagnetic field lines. The latter property implies that the backscatter cross-section is maximum when the radar is pointing almost perpendicular to the field line, although there have been several instances of auroral backscatter occurring at large aspect angles.

Other physical mechanisms for producing the auroral irregularities have also been proposed. A critical review of plasma irregularities in the auroral electrojet, covering both theory and experiment, has been given by Sahr and Fejer (1996).

Polarization

Investigations of the polarization of backscatter from auroral E-region irregularities have concluded that coherent scatter of spectral classes 1, 2, and 3 has similar polarization characteristics. For most of the observations, the scattering of a linearly polarized incident wave produced an essentially linear and highly polarized scattered wave, implying that there was a small scattering volume and/or a small number of discrete scatterers located close to one another. This also confirms that the scattering process is a weak coherent one. There were, however, some significant exceptions.

Observing geometry and occurrence

The region which may be observed is restricted by the aspect sensitivity. As Figure 6.31 illustrates, the radar must be situated at middle latitude if the beam is to intersect field-aligned irregularities in the auroral zone.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

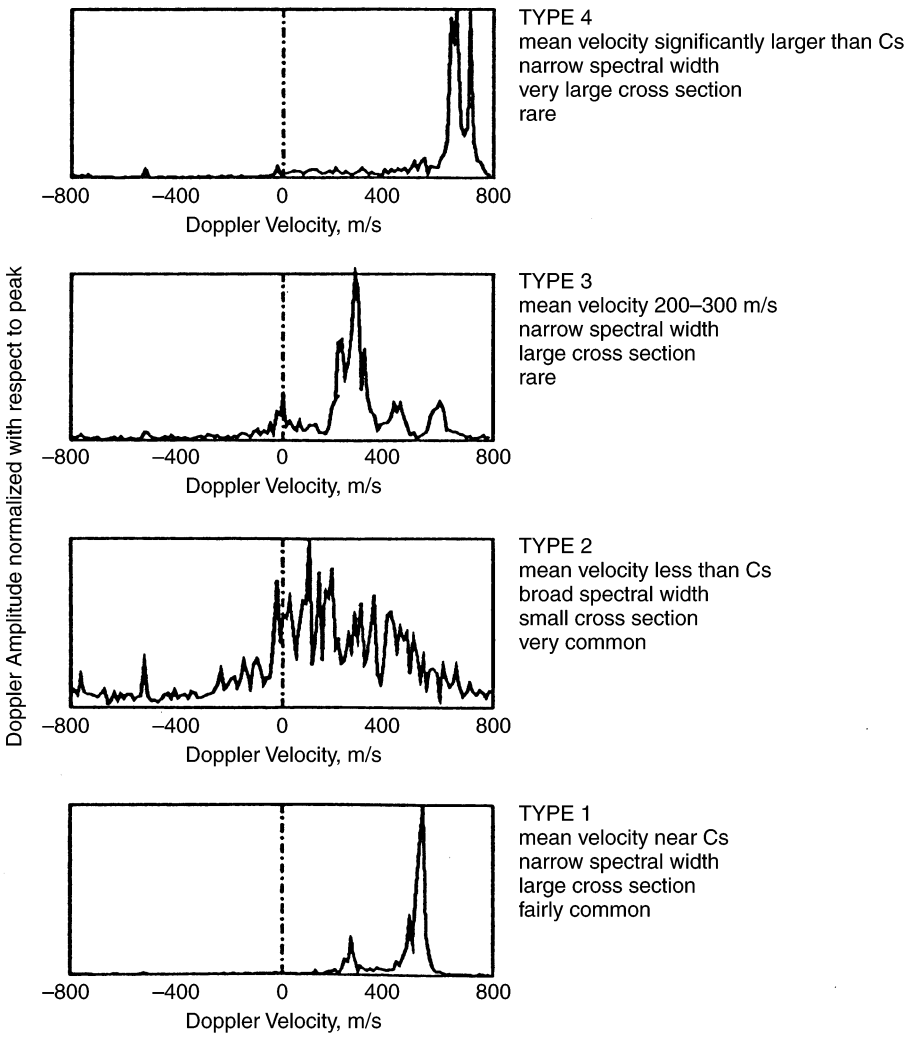


Figure 6.30. Four types of echo in auroral radar. The observations were with a 50-MHz radar in March 1989, and each analysis was based on 20 s of data. Cs is the local speed of sound. (J. D. Sahr and B. G. Fejer, *J. Geophys. Res.* **101**, 26893, 1996, copyright by the American Geophysical Union.)

Broadly speaking, the diurnal and seasonal occurrence of the radio aurora is similar to that of the visual aurora, except, of course, during daylight when the visual aurora cannot be seen. The strongest echoes occur near 65° latitude, and during disturbances the echoing region extends equatorward. The echoes can be detected at any time of day, but are most pronounced near local magnetic midnight.

The phenomenon of E-region irregularities is closely related to the behavior of the auroral electrojet. Sahr and Fejer (1996) draw attention to the problem of modeling them globally, to their importance in radio propagation, and to gaps in methods of data analysis.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

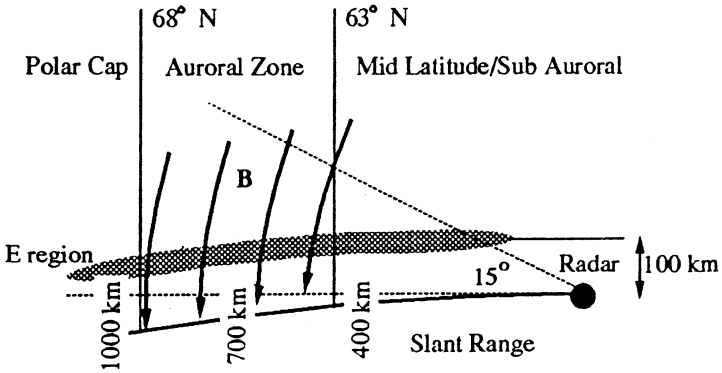


Figure 6.31. The geometry of coherent radar-scatter at high latitude. For significant scattering the beam must be normal to the field-aligned irregularites to within about 2°. In the absence of refraction (at VHF and UHF) the range will be between 400 and 1200 km. (J. D. Sahr and B. G. Fejer, *J. Geophys. Res.* **101**, 26893, 1996, copyright by the American Geophysical Union.)

6.5.6 Auroral infrasonic waves

Coming within the acoustic-wave domain of atmospheric waves (Section 1.6), *auroral infrasonic waves* (*AIWs*) are an interesting, though not very well known, feature in the spectrum of high-latitude phenomena (Wilson, 1969). They originate in the supersonic horizontal motion of the large-scale electrojets that flow within auroral arcs. The motion produces an infrasonic "bow wave," whose propagation is highly anisotropic and which reaches the Earth's surface about 6 min after the zenith passage of the arc. AIWs can sometimes be detected at the ground as much as 1000 km from the source. The period will be in the range 10–100 s – a frequency range of 0.01–0.1 Hz – with maximum spectral power density at a period of about 70 s, the pressure wave having amplitude from 0.5 to 20 μbars. They are detected with microphone arrays and are found to be highly coherent across arrays whose sensors are spaced by up to 6 km. The speed at which an AIW crosses the array will be between about 300 and 1000 m s^{-1}, the average being about 500 m s^{-1} (Wilson *et al.*, 1976). An example observed with a microphone array at Fairbanks, Alaska, is shown in Figure 6.32.

AIWs occur on the night side of the Earth, having a diurnal occurrence maximum near local midnight and a seasonal maximum around the equinoxes. Episodes of AIW activity, *AIW substorms*, are highly correlated to the occurrence of negative magnetic bays (Section 2.5.3). The horizontal component of the AIW's velocity is parallel to the motion of the supersonic auroral arc and also parallel to the horizontal magnetic-disturbance vector associated with the westward electrojet within the arc.

It is an interesting fact that AIWs are generated only by arcs moving equatorward; those moving poleward have no such effect. It has also been noted (Wilson and Hargreaves, 1974) that, statistically speaking, their direction of movement is similar to that of peaks in auroral radio absorption (Section 7.2.4) at a similar lat-

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

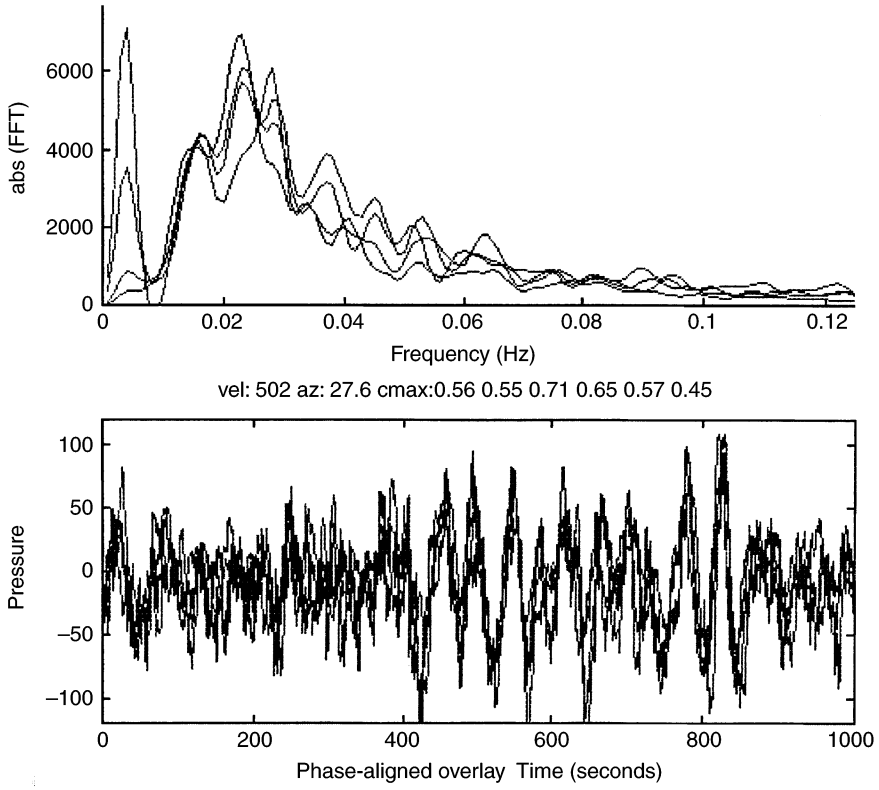


Figure 6.32. An auroral infrasonic wave observed with a detector array at Fairbanks, Alaska. The upper panel shows the power spectrum of the event, and the lower one shows the waveform. The wave traveled at 502 m s^{-1}, arriving from azimuth 27.6°. (R. D. Hunsucker, private communication.)

itude, having, in addition to the equatorward motion, a westward component before midnight and an eastward one after midnight.

AIWs should be considered as part of the total energy budget of terrestrial auroral phenomena and as a sensitive sensor of dynamic auroral arcs.

6.5.7 The generation of acoustic gravity waves

Another consequence of the electron-precipitation events, which create large electron densities supporting electrojets, is that the auroral E region is a source of *acoustic-gravity waves (AGWs)*. Strictly, there are two mechanisms. One is intense Joule heating ($\mathbf{J} \cdot \mathbf{E}$), where $\mathbf{J}$ is the current density and $\mathbf{E}$ is the electric field, which occurs in localized regions. The other is the Lorenz force ($\mathbf{J} \times \mathbf{B}$), where $\mathbf{B}$ is the flux density of the geomagnetic field. AGWs were introduced in Section 1.6 and it has been shown that those in the "large-scale" category originate in the auroral regions, probably from one of these sources. In the ionosphere the AGW is recognized as a *traveling ionospheric disturbance (TID)*, which propagates in the F region, primarily equatorward, for distances that may exceed 10000 km.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

In an investigation of TIDs in electron content at $L = 4$, most of which were in the "medium-scale" range with period 20–60 min, Hunsucker and Hargreaves (1988) noted that the waves were present almost continuously during daylight hours at the level of 1–4% of the electron content. Although no specifc source was identified, it must be significant that the incidence of these waves was far greater at $L = 4$ than it is at middle latitudes.

Some of the energy deposited in the auroral ionosphere from the magnetosphere may be transported to other latitudes by the action of AGWs, as well as by neutral-air winds and tides. It is estimated that 5–10% of this redistribution is due to AGWs.

6.6 Summary and implications

Except for the very large seasonal variability, the polar E region is relatively benign, compared with the auroral region. Precipitation of 1–10 keV electrons along geomagnetic field lines through the magnetospheric plasma sheet into the auroral ionosphere produces several very important effects: the luminous aurora, anomalously high E-region electron densities (conductivity), and localized regions of intense Joule heating and Lorenz forces. These phenomena are organized by the geomagnetic field into the northern and southern auroral ovals, which are stationary in space in Sun–Earth coordinates, with the Earth rotating underneath. The ovals are centered at approximately 67° geomagnetic latitude near magnetic midnight and 77° geomagnetic latitude near geomagnetic noon under "quiet conditions," and the latitudinal "thickness" of the oval increases as K_p increases. The most used auroral oval models are those derived from visual auroral observations, which give a reasonable estimate of auroral-E ionization for K_p values up through $K_p = 7$. Other ovals based on the TIROS and DMSP satellite particle measurements, which may give a more accurate mapping both of the electron precipitation which produces the auroral-E ionization and of the particle precipitation which produces D-region absorption and F-region irregularities, are also available.

The altitude profiles of auroral luminosity and of electron density in the E region are almost identical in shape. Electron densities as high as 4.4×10^6 electrons cm^{-3} have been measured by the Chatanika ISR during a large geomagnetic storm, and densities from $\sim 5 \times 10^5$ to 10^6 electrons cm^{-3} are quite common around magnetic midnight near 65° north geomagnetic latitude (Fairbanks, Alaska). The College (Fairbanks) ionosonde has also measured an auroral-E top frequency of 13 MHz, which corresponds to an oblique frequency of 57 MHz on a 1000-km long, curved-Earth-limited one-hop propagation path. There is some question, however, regarding whether the auroral-E top frequency measured by an ionosonde is really a true plasma frequency.

The temporal and spatial behavior of the auroral-E layer is dynamic and is probably best demonstrated by observing the visual aurora and realizing that the

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

regions of highest intensity (brightness) are, indeed, regions of high auroral-E electron density

The regions of high intensity in the visual aurora are also regions of intensified conductivity (hence electric current) of the auroral electrojet in the E region. These enhanced currents can produce intense Joule heating and Lorenz forces, which in turn generate atmospheric gravity waves (AGWs) that couple with the electron plasma to produce traveling ionospheric disturbances (TIDs). The large-scale TIDs may travel equatorward in the F region for over 10000 km, creating anomalous propagation in mid-latitude HF systems.

Another effect of auroral-oval E-region dynamics is the generation of auroral infrasonic waves (AIWs), which occur when an auroral arc travels supersonically towards the equator. For certain conditions of electron density, auroral brightness, particle-precipitation energies, Mach number, and orientation of the arc with respect to the geomagnetic field, a "bow shock wave" is formed, launching AIWs that may then be detected by a suitable array of acoustic sensors on the ground.

6.7 References and bibliography

6.2 Statistical distribution of the aurora

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

Akasofu, S.-I. (1974) The aurora and the magnetosphere; the Chapman memorial lecture. *Planet. Space Sci.*, **22**, 885.

Chubb, T. A. and Hicks, G. T. (1970) Observations of the aurora in the far ultraviolet from OGO 4. *J. Geophys. Res.* **75**, 1290.

Feldstein, Y. I. and Starkov, G. V. (1967) Dynamics of auroral belt and polar geomagnetic disturbances. *Planet. Space Sci.* **15**, 209.

Frank, L. A. and Craven, J. D. (1988) Imaging results from Dynamics Explorer 1. *J. Geophys. Res.* **26**, 246.

Fuller-Rowell, T. J. and Evans, D. S. (1987) Height-integrated Pedersen and Hall conductivity patterns inferred from TIROS-NOAA satellite data. *J. Geophys. Res.* **92**, 7606.

Goodman, J. M. (1992) *HF Communications – Science and Technology*. Van Nostrand Reinhold, New York.

Gussenhoven, M. S., Hardy, D. A., and Heinemann, N. (1983) Systematics of the equatorward diffuse auroral boundary. *J. Geophys. Res.* **88**, 5692.

Hardy, D. A., Gussenhoven, M. S., and Holeman, E. (1985) A statistical model of auroral electron precipitation. *J. Geophys. Res.* **90**, 4229.

Hardy, D.A., Gussenhoven, M. S., and Brautigan, D. (1989) A statistical model of auroral ion precipitation. *J. Geophys. Res.* **94**, 370.

Meng, C.-I. (1984) Dynamic variation of the auroral oval during intense magnetic storms. *J. Geophys. Res.* **89**, 227.

Meng, C.-I. and Makita. K. (1986) Dynamic variations of the polar cap. *Solar Wind–Magnetosphere Coupling* (eds. Kamide and Slavin), p. 605. Terra Scientific, Tokyo.

Vestine, E. H. (1944) The geographic incidence of aurora and magnetic disturbance, Northern Hemisphere. *Terr. Magn. Atmos. Electricity* **49**, 77.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

Whalen, J. A. (1970) Auroral oval plotter and nomograph for determining geomagnetic local time, latitude and longitude in the Northern Hemisphere. Report AFCRL-70-0422, Environmental Research Paper 327. (From Defense Technical Information Center, Cameron Station, Alexandria, VA 22314, USA)

6.3 The auroral phenomena

Bauer, S. J. (1973) *Physics of Planetary Ionospheres*. Springer-Verlag, Berlin.

Gazey, N. G. J., Smith, P. N., Rijnbeek, R. P., Buchan, M., and Lockwood, M. (1996) The motion of auroral arcs within the convective plasma flow. *Third International Conference on Substorms*, Versailles, France. Report ESA SP-389, p. 11.

Harang, L. (1951) *The Aurorae*. Wiley, New York.

Hargreaves, J. K. (1969) Auroral absorption of HF radio waves in the ionosphere – a review of results from the first decade of riometry. *Proc. Inst. Elect. Electronics Engineer* **57**, 1348.

Hartz, T. R. and Brice, N. M. (1967) The general pattern of auroral particle precipitation. *Planet. Space Sci.* **15**, 301.

Rees, M. H. (1989) *Physics and Chemistry of the Upper Atmosphere*. Cambridge University Press, Cambridge.

Störmer, C. (1955) *The Polar Aurora*. Oxford University Press, Oxford.

6.4 The substorm

Akasofu, S.-I. (1964) The development of the auroral substorm. *Planet. Space Sci.* **12**, 273.

Akasofu, S.-I. (1968) *Polar and Magnetospheric Substorms*. Springer-Verlag, New York.

Akasofu, S.-I. (1970) In memoriam Sydney Chapman. *Space Sci. Rev.* **11**, 599.

Akasofu, S.-I. (1974) A study of auroral displays photographed from the DMSP-2 satellite and from the Alaska meridian chain of stations. *Space Sci. Rev.* **16**, 617.

Akasofu, S.-I. (1977) *Physics of Magnetospheric Substorms*. Reidel, Dordrecht.

Akasofu, S.-I. (1979) Interplanetary energy flux associated with magnetospheric storms. *Planet. Space Sci.* **27**, 425.

Akasofu, S.-I. and Chapman, S. (1961) The ring current, geomagnetic disturbance and the Van Allen radiation belts. *J. Geophys. Res.* **66**, 1321.

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

Angelopoulos, V. (1996) The role of impulsive particle acceleration in magnetotail circulation. *Third International Conference on Substorms*, Versailles, France. Report ESA SP-389, p. 17.

Birkeland, K. (1908) *The Norwegian Aurora Polaris Expedition 1902–3*, vol. 1, section 1. H. Aschehoug, Christiana.

Baumjohann, W. (1996) Storm–substorm relationship. *Third International Conference on Substorms*, Versailles, France. Report ESA SP-389, p. 627.

Borovsky, J. E., Nemzek, R. J., and Belian, R. D. (1993) The occurrence rate of magnetospheric-substorm onsets. *J. Geophys. Res.* **98**, 3807.

Clauer, C. R. and McPherron, R. L. (1980) The relative importance of the interplane-

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

tary electric field and magnetospheric substorms on partial current development. *J. Geophys. Res.* **85**, 6747.

Elphinstone, R. D., Murphree, J. S., Cogger, L. L., Hearn, D., and Henderson, M. G. (1991) Observations of changes to the auroral distribution prior to substorm onset. *Magnetospheric Substorms* (eds. J. R. Kan, T. A. Potemra, S. Kokubun, and T. Iijima), p. 257. American Geophysical Union, Washington DC.

Elphinstone, R. D., Murphree, J. S., and Cogger, L. L. (1996) What is a global auroral substorm? *Rev. Geophys.* **34**, 169.

Feldstein, Y. I. (1991) Substorm current systems and auroral dynamics. *Magnetospheric Substorms* (eds. J. R. Kan, T. A. Potemra, S. Kokubun and T. Iijima), p. 29. American Geophysical Union, Washington DC.

Hargreaves, J. K. (1996) Substorm effects in the D region. *Third International Conference on Substorms*, Versailles, France. Report ESA SP-389, p. 663.

Kamide, Y. (1991) The auroral electrojets: relative importance of ionospheric conductivities and electric fields. *Auroral Physics* (eds. C.-I. Meng, M. J. Rycroft, and L. A. Frank), p. 385. Cambridge University Press, Cambridge.

Lui, A. T. Y. (1991) Extended consideration of a synthesis model for magnetospheric substorms. *Magnetospheric Substorms* (eds. J. R. Kan, T. A. Potemra, S. Kokubun, and T. Iijima), p. 43. American Geophysical Union, Washington DC.

Lui, A. T. Y. (1992) Magnetospheric substorms. *Phys. Fluids* B, 4, 2257.

McPherron, R. L., Russell, C. T., and Aubry, M. P. (1973) Satellite studies of magnetospheric substorms on August 15, 1968: 9. Phenomenological model for substorms. *J. Geophys. Res.* **78**, 3131.

Murphree, J. S., Elphinstone, R. D., Cogger, L. L., and Hearn, D. (1991) Viking optical substorm signatures. *Magnetospheric Substorms* (eds. J. R. Kan, T. A. Potemra, S. Kokubun, and T. Iijima), p. 241. American Geophysical Union, Washington DC.

Pulkkinen, T. I. (1996) Pseudobreakup or substorm? *Third International Conference on Substorms*, Versailles, France. Report ESA SP-389, p. 285.

Rostocker, G. (1991) Some observational constraints for substorm models. *Magnetospheric Substorms* (eds. J. R. Kan, T. A. Potemra, S. Kokubun, and T. Iijima), p. 61. American Geophysical Union, Washington DC.

Rostoker, G., Akasofu, S.-I., Foster, J., Greenwald, R. A., Kamide, Y., Kawasaki, K., Liu, A. T. Y,. McPherron, R. L., and Russell, C. T. (1980) Magnetospheric substorms – definitions and signatures. *J. Geophys. Res.* **85**, 1663.

Shiokawa, K., Baumjohann, W., and Haerendel, G. (1997) Braking of high-speed flows in the near-Earth tail. *Geophys. Res. Lett.* **24**, 1179.

Stern, D. P. (1991) The beginning of substorm research. *Magnetospheric Substorms* (eds. J. R. Kan, T. A. Potemra, S. Kokubun, and T. Iijima), p. 11. American Geophysical Union, Washington DC.

Stern, D. P. (1996) A brief history of magnetospheric physics during the space age. *Rev. Geophysics* **34**, 1.

6.5 The E region at high latitude

Bates, H. F. and Hunsucker, R. D. (1974) Quiet and disturbed electron density profiles in the auroral zone ionosphere. *Radio Sci.* **9**, 455.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016

Hunsucker, R. D. (1975) Chatanika radar investigation of high-latitude E-region ionization structure and dynamics. *Radio Sci.* **10**, 277.

Hunsucker, R. D. and Hargreaves, J. K. (1988) A study of gravity waves in ionospheric electron content at $L = 4$. *J. Atmos. Terr. Phys.* **50**, 167.

Rees, M. H. (1989) *Physics and Chemistry of the Upper Atmosphere*. Cambridge University Press, Cambridge.

Sahr, J. D. and Fejer, B. G. (1996)Auroral electrojet plasma irregularity theory and experiment: a critical review of present understanding and future directions. *J. Geophys. Res.* **101**, 26 893.

Wilson, C. R. (1969) Auroral infrasonic waves. *J. Geophys. Res.* **74**, 1812.

Wilson, C. R. and Hargreaves, J. K. (1974) The motions of peaks in ionospheric absorption and infrasonic waves. *J. Atmos. Terr. Phys*. **36**, 1555.

Wilson, C. R., Hunsucker, R. D., and Romick, G. J. (1976) An auroral substorm investigation using Chatanika radar and other geophysical sensors. *Planet. Space Sci.* **24**, 1155.

General reading on the aurora and related topics

Books

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

Akasofu, S.-I. (1977) *Physics of Magnetospheric Substorms.* Reidel, Dordrecht.

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

Brekke, A. (1997) *Physics of the Upper Polar Atmosphere*. Wiley, Chichester.

Brekke, A. and Egeland, A. (1983) *The Northern Lights – From Mythology to Space Research*. Springer-Verlag, Berlin.

Chamberlain, J. W. (1961) *Physics of the Aurora and Airglow.* Academic Press, New York.

Eather, R. H. (1980) *Majestic Lights*. American Geophysical Union, Washington, DC.

Kamide, Y. and Baumjohann, W. (1993) *Magnetosphere–Ionosphere Coupling*. Springer-Verlag, Berlin.

Kan, J. R., Potemra, T. A., Kokubun, S., and Iijima, T. (eds.) (1991) *Magnetospheric Substorms*. American Geophysical Union, Washington,DC.

Kennel. C. F. (1995) *Convection and Substorms*. Oxford University Press, Oxford.

Omholt, A. (1971) *The Optical Aurora*. Springer-Verlag, Berlin.

Störmer, C. (1955) *The Polar Aurora*. Oxford University Press, Oxford.

Vallance Jones, A. (1974) *Aurora*. Reidel, Dordrecht.

Conference reports

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

McCormac, B. M. (ed.) (1967) *Aurora and Airglow*. Van Nostrand Reinhold Co., New York.

McCormac, B. M. and Omholt, A. (eds.) (1969) *Atmospheric Emissions*. Van Nostrand Reinhold Co., New York.

McCormac, B. M. (ed.) (1971) *The Radiating Atmosphere*. Reidel, Dordrecht.

Meng, C.-I., Rycroft, M. J., and Frank, L. A. (eds.) (1991) *Auroral Physics*. Cambridge University Press, Cambridge.

Downloaded from Cambridge Books Online by IP 141.211.4.224 on Wed Mar 23 15:42:29 GMT 2016.
http://dx.doi.org/10.1017/CBO9780511535758.008
Cambridge Books Online © Cambridge University Press, 2016