

Det Norske Videnskaps-Akademi i Oslo

Geofysiska Notiser

GEOFYSISKE PUBLIKASJONER

VOL. 30 NO. 4 1973

UNIVERSITETSFÖRLAGET

Geophysica Norvegica

is a journal of geophysics, issued under the auspices of the Norwegian Academy of Science and Letters in Oslo

EDITOR

Eigil Hesstvedt, Institutt for geofysikk, Universitetet i Oslo, Oslo 3, Norway.

EDITORIAL BOARD

Olaf Devik, Rektorhaugen 11, Oslo, Norway

Olav Holt, Nordlysobservatoriet, Universitetet i Tromsø, 9000 Tromsø, Norway

Håkon Mosby, Geofysisk institutt, Universitetet i Bergen, 5000 Bergen, Norway.

PUBLISHER

Universitetsforlaget: P. O. Box 307, Blindern, Oslo 3, Norway.

P. O. Box 142, Boston, Mass. 02113, U.S.A.

SUBSCRIPTION

Geophysica Norvegica (Geofysiske Publikasjoner) is published at irregular intervals. Order from the Publisher, Universitetsforlaget.

Geophysica Norvegica (Geofysiske Publikasjoner) is a series of scientific publications issued by the Norwegian Academy of Science and Letters in Oslo. The Geophysical Commission appoints an editor and editorial committee.

Manuscripts for publication in *Geophysica Norvegica* should be carefully prepared (see Instructions to Author on p. 3 of the cover) and sent to the editor. The next step is the manuscript's submission to the Academy by a competent member, who is responsible for ensuring that the paper meets a sufficiently high scientific standard. (Members of the Academy are allowed to submit their own papers.) The final decision as to whether the paper should be published is taken by the editor.

Geophysica Norvegica is mainly intended as a journal for Norwegian authors, but papers from other authors may be accepted provided that the work has been carried out at a Norwegian institution or its content has a special relevance to Norway.

Electric Field Microstructures in the Auroral E-Region

J. A. HOLTET

The Norwegian Institute of Cosmic Physics, Blindern, Oslo, Norway

Holtet, J. A. Electric Field Microstructures in the Auroral *E*-Region. *Geophysica Norvegica*, Vol. 30, No. 4, 1973.

The present work represents a study of microstructures in the auroral *E*-region based on in situ observations of AC electric fields. The main characteristics of the observed wavefields, called ELF *E*-layer noise, are as follows:

- 1) The emissions appear in the altitude region 95–120 km.
- 2) The noise is found in the range 50 Hz to 3 kHz.
- 3) The orientation in space of the wave *E*-field vector is well defined and is approximately in the magnetic east-west direction. This does, however, change significantly with altitude.
- 4) The wave is electrostatic.

It is shown that these fields can be generated by a two-stream plasma instability. The observed characteristics are compared with features derived using linear instability theory and an ionospheric model representing the observation conditions, and a good correspondence is found. An additional piece of information obtained from the measurements is that the wavefield observations give means to record small scale variations in the direction of the electric current.

J. A. Holtet, The Norwegian Institute of Cosmic Physics, Blindern, Oslo 3, Norway

The day will yet come when the progress of research through long ages will reveal to sight the mysteries of nature that are now concealed. A single lifetime, though it were wholly devoted to study, does not suffice for the investigation of problems of such complexity... It must, therefore, require long successive ages to unfold all. The day will yet come when posterity will be amazed that we remained ignorant of things that will to them seem so plain.

Seneca. *Naturales quaestiones* 7.25 (about A.D. 63).

CONTENTS

	Pages
1. INTRODUCTION	4
2. IN SITU MEASUREMENTS OF <i>E</i> -REGION ELECTRIC FIELD MICRO- STRUCTURES.....	4
2.1. Introduction	4
2.2. Launch conditions and payload operation	5
2.3. AC measurements above 130 km	8
2.4. AC measurements below 130 km	10
2.5. Other observations of ELF <i>E</i> -layer noise	13
2.6. Discussion of the raw data	13
3. CHARACTERISTICS OF THE IONOSPHERIC MEDIUM	14
3.1. Introduction	14
3.2. The neutral atmosphere and its properties	14
3.3. Production and loss of ionization	16
3.4. Ionization by precipitating particles	20
3.5. Ionospheric parameters	21
4. ELECTRODYNAMICS OF THE IONOSPHERE	25
4.1. Introduction	25
4.2. Particle motion and conductivities	25
4.3. Drift motion of ionization irregularities.....	27
5. CURRENT SYSTEMS AND LARGE-SCALE ELECTRIC FIELDS IN THE HIGH LATITUDE IONOSPHERE.....	29
5.1. Introduction	29
5.2. Ionospheric current systems	29
5.3. Sources of DC electric fields	33
5.4. Characteristics of large-scale electric fields	34
5.5. Experimental techniques for in situ measurements of electric fields	37
5.6. Event studies of DC electric fields	
6. VLF DOPPLER OBSERVATIONS OF IONIZATION MICROSTRUCTURES	49
6.1. Introduction	49
6.2. The Doppler technique; theoretical background	49
6.3. Recording and analyzing instrumentation	51
6.4. Microstructures in the electron density distribution.....	52
7. STRUCTURES IN THE AURORAL <i>E</i> -REGION	54
7.1. Introduction	54
7.2. Small-scale structures in optical aurora	54
7.3. Finestructures in particle precipitation	55
7.4. Radio aurora	56
8. PLASMA INSTABILITIES AND <i>E</i> -REGION IRREGULARITIES	59
8.1. Introduction	59
8.2. The crossed fields instability	59
8.3. Ion wave instabilities – The two-stream instability	60

9. DISCUSSIONS AND CONCLUSIONS	68
9.1. An ionospheric model and the related two-stream instability	68
9.2. Amplitude <i>vs.</i> frequency distribution of ELF <i>E</i> -layer noise	72
9.3. Direction variation of the wave electric field and corresponding changes in current distribution	73
9.4. Conclusions.....	74
APPENDIX I: VARIATIONS IN PROBE POTENTIAL CAUSED BY ENVIRONMENTAL EFFECTS	75
AI.1. Probe motion, velocity gradients and their influence on the floating potential	75
AI.2. Error analysis.....	76
APPENDIX II: THE IMPEDANCE OF A CYLINDRICAL PROBE IN A PLASMA	82
ACKNOWLEDGEMENTS	84
REFERENCES	85

1. INTRODUCTION AND OUTLINE

The main purpose of this study was to perform an experimental and theoretical investigation of small-scale irregularities in the auroral *E*-region. 'Small-scale' in this context means structures with horizontal extent in the range 10^1 – 10^{-1} m.

The experimental basis was in situ observations of AC electric fields during the presence of visual aurora. Variational fields of this kind can be directly related to irregularities in the distribution of thermal electrons. Their presence has earlier been postulated from measurements of other ionospheric parameters and on theoretical grounds. The technique utilized in the present study has, however, provided a new approach to the problems. In situ observations give first-hand information about spectral distribution and direction of the wavefield, together with variations of the wave properties *vs.* altitude. The direct result of this is improved mapping of the small-scale electric fields in the *E*-region. This, in turn, leads to a better understanding of physical processes involved in the formation of irregularities in this part of the ionosphere.

To give an impression of the kind of irregularities which will be considered, and outline the problem, we will begin with a presentation of the experimental results together with a short discussion of the raw data (Chapter 2). In dealing with ionospheric irregularities it is, however, necessary to obtain information about the 'regular' ionosphere. In Chapter 3, therefore, we rather summarily give a picture of the ionospheric medium, and present various ionospheric parameters and their variations, while Chapter 4 deals with the electrodynamics of the ionosphere – the motion of charged particles influenced by electric and magnetic fields. This is followed in Chapter 5 by a presentation of auroral current systems, a result of the particle motions. In the same chapter we will evaluate the present knowledge about large-scale electric fields, the driving force of the currents. Furthermore, we will present some observations of DC fields in auroral situations and go into the experimental technique for measuring electric fields in the ionosphere. Details of this subject will be further discussed in Appendices I and II.

Supporting the observations of small-scale electric fields, presented in Chapter 2, is the presence of microstructures in the electron distribution observed at the same time as the field variations. These recordings, made by the VLF Doppler technique, are described in Chapter 6. To see our observations in a wider perspective, the subject for the following chapter is small-scale irregularities in the auroral *E*-region in general, giving a survey of observed small-scale structures in particle precipitation, optical aurora, and radio aurora.

The source of *E*-region microstructures is likely to be found in a plasma instability (Chapter 8). It is deduced that the observed features during the existing ionospheric conditions can plausibly be explained by a modified two-stream instability.

The last chapter is devoted to comparisons and discussions of the instability theory *vs.* observed features. The primary parameter in generation of the instability is the electron-ion drift velocity. Both frequency, amplitude, and direction of the wavefield depend on this property. A supplementary result from the observations is therefore that from observations of the field direction one can track the velocity vector and thus record small-scale variations in the direction of the current.

The observation of the *E*-region irregularities came as an unexpected result from an experiment originally flown to measure very low frequency radio emissions in aurora. The payload was therefore not primarily composed to give thorough information about the *E*-region plasma. In the comparison between observed and theoretically derived properties of the wavefield one is therefore limited to the use of a somewhat simplified ionospheric model. However, essential observed features as altitude region, amplitude *vs.* frequency distribution, and direction of the wavefield in space, correspond very well to that predicted from the theoretical model study.

2. IN SITU MEASUREMENTS OF *E*-REGION ELECTRIC FIELD MICROSTRUCTURES

2.1. Introduction

The observations which constitute the basis for this study, and which will later be put into a wider

perspective, are the results from measurements of small-scale variations in the electric field. The observations were made by means of rockets, and are based on the double floating probe technique. The measuring technique and instrumentation will be described in detail in Chapter 5 and Appendix I.

The data mainly originate from two rocket flights, F23 and F24, but, for comparison, results from two other rockets (NASA 14.383 and 18.57) will also be mentioned. The main point will be the presence of small-scale fields in the *E*-region, but to get an impression of the entire picture, the AC observations at other altitudes will also be briefly described.

2.2 Launch conditions and payload operation

The rockets were launched from Andøya Rocket Range in Northern Norway. Coordinates for the range are: Geographic 69.3°N, 16.0°E, corrected geomagnetic 66.1°N, 103.1°E (Hakura 1965). Magnetic midnight (eccentric dipole time) is about 2200 UT. The launch site is located approximately 120 km magnetic west of Tromsø (Auroral Observatory), 230 km magnetic north-west of Kiruna (Rocket Range and Geophysical Observatory), and 600 km magnetic south-west of Bear Island (magnetic observatory).

The rockets were night flights, and were fired during weakly to moderately disturbed conditions. Visual aurora was present in or close to the launch direction of F23, F24, and NASA 18.57, while no aurora could be seen near the trajectory of NASA 14.383. The main characteristics of the payloads, together with information about ionospheric disturbances, are summarized in Table 2.1. The launch conditions for F24 and F23 will be described in more detail.

F24: Even though ground observations of the ionosphere at the launch time of F24 (0017.37 UT, March 8, 1970) gave evidence of moderate activity, the period when the launch took place must be described as very disturbed. A SC at 0805 UT, March 5 preceded a minor storm and a second SC, which initiated a great storm, started at 1420 UT, March 8. However, the conditions during the launch period were dominated by a minor PCA which was reported to start about 1410 UT, March 6. The ATS-1 records indicated a maximum for the proton event some time before 0020 UT, March 8 (ESSA 1970). A PCA caused by precipitation of high energy solar particles will establish relatively stable ionospheric conditions over a large scale (Hultqvist 1968). The main features will thus be conserved over a long

Table 2.1. Flight characteristics and launch conditions

Rocket code	F24	F23	NASA 14.383	NASA 18.57
Rocket type	Nike/Tomahawk	Nike/Tomahawk	Nike/Apache	Nike/Tomahawk
Launch date	March 8, 1970	Nov. 13, 1970	Jan. 12, 1972	Jan. 16, 1973
Launch time ($T = 0$), UT	0017.37	2219.41	2116.35	2058.51
Azimuth setting, °	320	320	343	345.5
Peak altitude, km	216	214	197	260
Spin period, <i>s</i>	1.8	1.09	0.48	1.1
Coning period, <i>s</i>	260	175	40	170
Coning half angle, °	20	17.5		
<i>E</i> -field exp.	DC, widebd., spectr.	DC, widebd., spectr.	DC, spectr.	DC, spectr.
Auroral cond.	Glow	Patchy glow	No near trajectory	Glow, double arc
<i>K_p</i>	4+	2o	2-	3-
<i>D_{st}</i> , γ	-43	-18		
CNA(27,6 MHz, $T = 0$), dB	0.6	0.7	0.2	4.5
Magn. act. ($T = 0$), γ	-300 (ΔH)	-240 (ΔH)	+50 (ΔX)	-150 (ΔX)

period, but this does not prevent other perturbations from being superimposed on the slow PCA variation.

The launch took place in the recovery of a negative bay, and the deflection in the H -component decreased fairly steadily from -290 to -190γ during the flight (cf. Fig. 2.1, curve b). The deviation in magnetic H and Z at Bear Island, Tromsø, and Kiruna, gave evidence of a westward electrojet with the integrated current system located to the north of, but close to the launch site zenith. Thus, the rocket trajectory was most likely to the north of the main electrojet. The K_p value of 4 (6- in the previous 3-hour interval) and the Dst figure of -43γ showed a high magnetic activity.

The auroral situation was dominated by a fairly homogeneous and quiet glow which covered the launch direction. The intensity of the N_2^+

(4278) emission as measured from the ground decreased slowly from about 1.2 to 0.9 kR, without any marked structure (cf. Fig. 2.1.c). Photographs indicated, however, the presence of very faint discrete forms within the glow.

Propagation studies of VLF waves, both from transmitters and naturally emitted (atmospherics), showed stable 'day time' propagation conditions, as are normal for a PCA event. This means that the precipitation of energetic particles caused ionization below 80 km resembling a day time D -layer (cf. Sect. 3.4). The cosmic noise absorption (27.6 MHz) at Tromsø, Andøya, and Kiruna, which gives an indication of the precipitation of high energy particles ($E > 40$ keV), was low (0.5 to 1 dB) but fairly constant (Fig. 2.1.a).

F23: This rocket was launched in a more typical auroral situation, with more structures and rapid

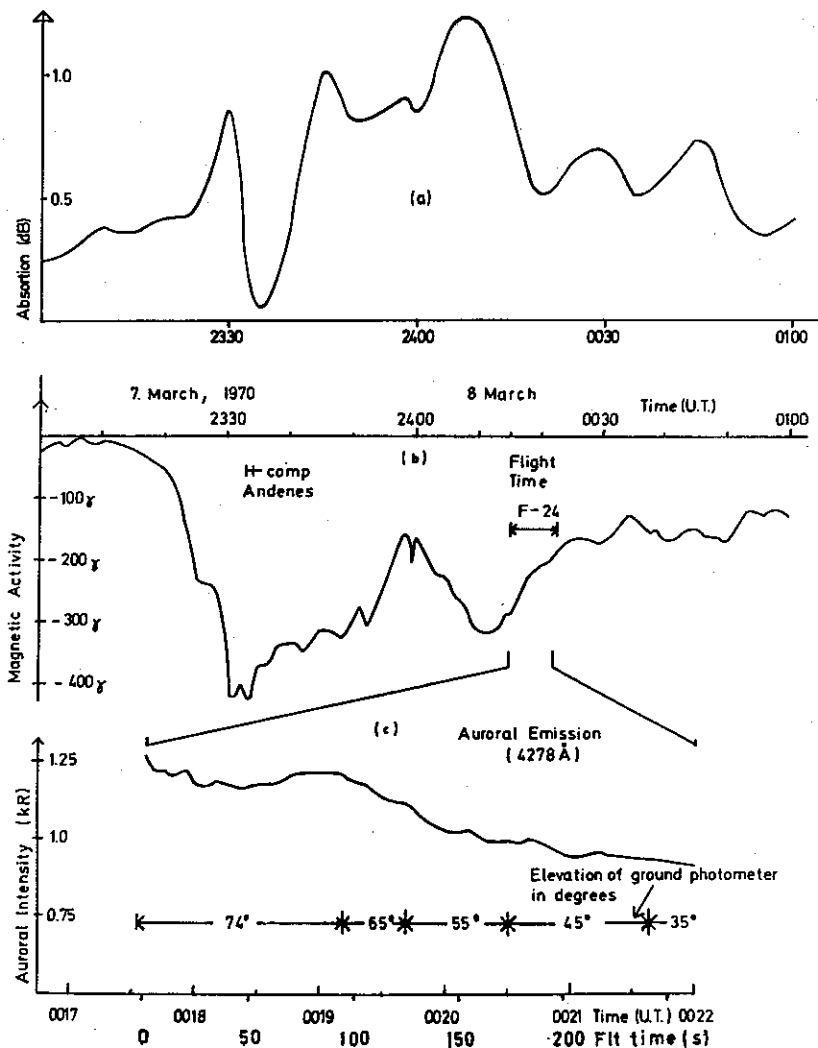


Fig. 2.1. Cosmic noise absorption at 27.6 MHz (curve (a)) and deviation in the magnetic H -component (curve (b)) measured at the launch site for the F24 launch period. Curve (c) shows the intensity of the auroral N_2^+ (4278 Å) emission as measured from the launch site at different elevation angles in the plane of the rocket trajectory. The elevation angles were selected to watch the particular parts of the aurora which were expected to be magnetically linked to the rocket during the different periods of the flight. The photometers were in a fixed position in each period.

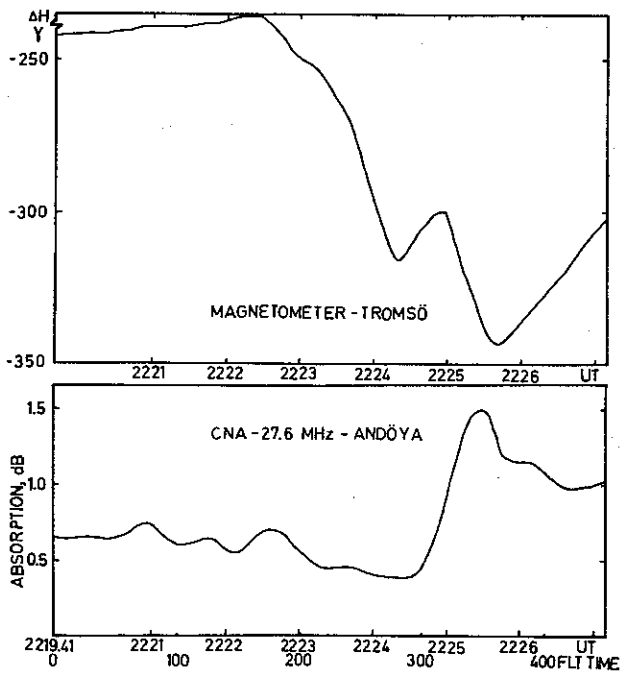


Fig. 2.2. Deviation in magnetic H -component (Tromsø) and cosmic noise absorption at 27.6 MHz (Andøya) during the flight of F23.

variations. Generally, the ionospheric conditions during the flight period can be characterized as moderately disturbed. The launch took place in the expansive phase of a substorm (*cf.* Sect. 5.2), but the situation as a whole was too complicated to allow a schematic classification. The deviation in magnetic H -component was at launch approximately -250γ , but during the flight this increased rather suddenly and reached a peak value of -350γ (*cf.* Fig. 2.2.a). The magnetic observations indicated a westward current system, consistent with the adopted current configuration for the substorm growth phase (*cf.* Sect. 5.2), with its centre somewhat to the north of the launch site. The K_p and Dst figures were respectively 20 and -18γ .

The cosmic noise absorption was also subject to large variations. From a fairly steady value of approximately 0.5 dB before and during the first part of the flight the absorption suddenly increased to about 1.5 dB, giving evidence of a sudden influx of energetic particles (Fig. 2.2.b). One may here find some resemblance with the F24 conditions, with energetic precipitation both in the PCA and in the sudden absorption event. However, the particle populations will most prob-

ably be quite different – while protons will dominate in the PCA precipitation, electrons will be most abundant in the SA flux (*cf. e.g.* Hultqvist 1968).

No photometric recordings of the aurora exist for this period, due to cloud coverage. From photographic observations at Tromsø the aurora can, however, be described as a patchy glow with weak structures near the launch direction. Vertical sounder records made at 2200 and 2300 UT at Kiruna and Tromsø showed the existence of sporadic E -layer. (In the case of F24 the low altitude absorption layer associated with the PCA precluded sounder measurements.)

Mechanically the two rockets were very similar, and in both cases the launcher settings were the same, with an azimuth angle of $\sim 320^\circ$. The peak altitudes were 216 km (F24) and 214 km (F23). F24 had a spin period after despin and antenna deployment of 1.8 s. The precession cone, as determined by two fluxgate aspect magnetometers, opened up to 20° half angle and the precession period was approximately 260 s. The orientation of the cone was such that the spin axis was parallel to \mathbf{B} once (near 204 km on the descent part). F23 had a spin period of 1.09 s, the axis of rotation was on a cone with 17.5° half angle, and the coning period was 175 s. In the attitude calculations it is assumed that the angular momentum vector lay in the plane of the trajectory.

In both payloads the E -field experiments functioned satisfactorily, without any problems of mechanical or electrical kind, and with a low interference level. The antennas were fully extended at 157 km for F24 and 138 km for F23. Atmospheric drag caused bending of the elements below 95 km on the descent path which disturbed the geometrical configuration and jammed the DC measurements (*cf.* Appendix I). AC data were obtained until the elements started to break off around 80 km. However, within the last 10 km a considerable amount of transient noise was present because of the unstable probe situation.

The DC measurements from the flights will be presented in Chapter 5. It will only be mentioned here that the general impression of the records fits the pictures of the launch conditions, with a

Discrete emissions, both risers and more complex forms, appeared sporadically throughout the whole flight in the frequency range 1 to 2.5 kHz. There was a maximum around 180 s, but this peak was not pronounced.

Several noise bands with different bandwidths (*cf.* Fig. 2.3) were present during most of the recording period between approximately 0.8 and 4.7 kHz. The intensity of the noise was partly spin modulated, showing an anisotropy in the wavefield. A broad, but pronounced maximum appeared between 150 and 230 s flight time (Fig. 2.4). This corresponds to a region with decreased DC field, and the fluctuations in DC and AC field strengths were also almost in antiphase. It should be noted that the spin modulation was less marked in this region.

Spin-modulated wideband noise (hiss) above 3 kHz occurred in two bursts, lasting from about 190 to 270 s and 280 to 375 s flight time. It should be noted that the minimum between 270 and 280 s corresponds to the time when the rocket axis was approximately parallel to the magnetic field lines.

The intensity of the signal integrated over one complete spin period was, however, rather slowly varying, and this emission showed no resemblance with the intense fluctuating emission known as auroral hiss (Helliwell 1965).

In general one can say that also the AC measurements reflect the stable and slowly varying PCA conditions.

The emission spectrum and variations recorded in the F23 event (summarized in Fig. 2.5) were basically different from that found in the F24 flight.

In this case hiss emissions were also present above approximately 6 kHz, but compared to the F24 hiss, these emissions were much more intensive and the amplitude fluctuated quite vigorously. According to VLF terminology these emissions should be classified as auroral hiss. Maximum signal strength appeared in several bursts between 210 and 250 s flight time. It will be noted that this corresponds to a peak in the DC field (*cf.* Sect. 5.6), and also that in the more rapid time variations the two records follow each other.

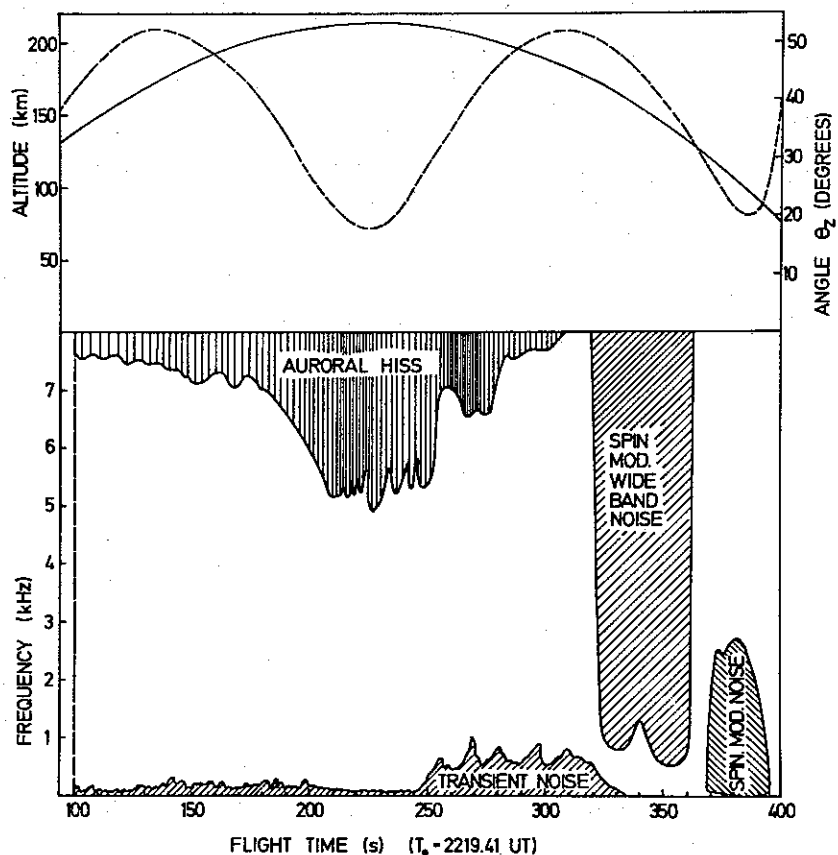
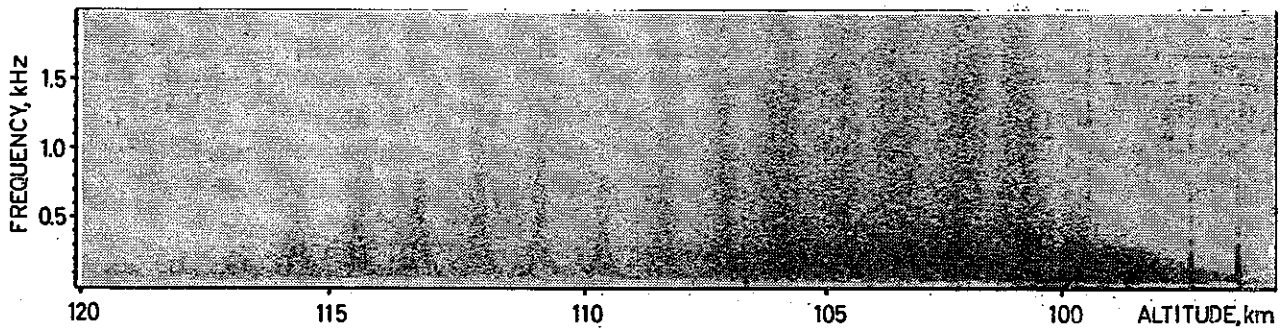
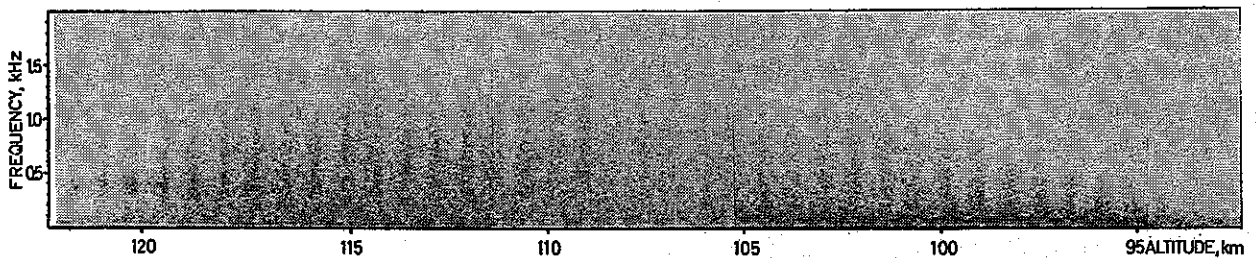


Fig. 2.5. Survey of emissions recorded during the flight of F23 together with a plot of altitude and angle between the rocket axis and the earth's magnetic field.

F 24, ANDÖYA, MARCH 8 1970

Fig. 2.6. Spectrogram of ELF *E*-layer noise recorded in the F24 event.

F 23, ANDÖYA, NOV. 13 1970

Fig. 2.7. Spectrogram of ELF *E*-layer noise in the F23 event.

This is quite contrary to the behaviour of the F24 noise bands.

In addition to the auroral hiss, two other kinds of emissions were observed. One was a band of noise bursts of rather transient character appearing below 1 kHz. The maxima in this noise appeared out of phase with the auroral hiss. The other type may be of more interest, as its onset can obviously be associated with the increased CNA, *i.e.* increased flux of high energy particles. The emission can be characterized as wideband noise from approximately 0.1 kHz to several tens of kHz. In the weaker parts it was spin modulated, but in the strongest bursts no spin modulation was present.

2.4. AC measurements below 130 km

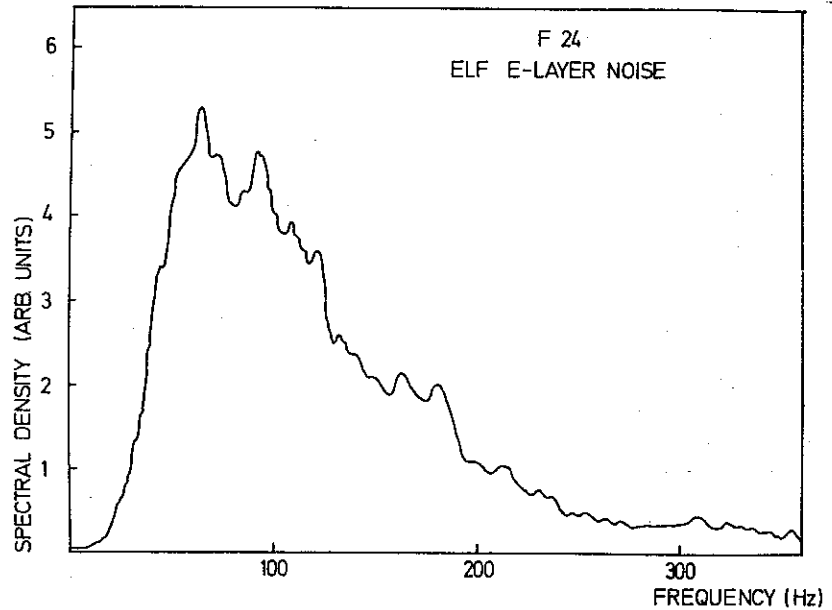
Both the launch conditions and the observations made during the flight were very different for these two rockets. Furthermore, the difference in coning angle and period made the attitude of the rocket body and *E*-field antenna with respect to the earth's magnetic field quite different, for the flight in general and below 130 km in particular. In spite of this the general appearance of the AC

electric fields observed in the height interval 130–90 km were almost identical for the two rockets. These fields, which spectrographically can be described as wideband noise with maximum intensity below 500 Hz (Figs. 2.6 and 2.7), will be termed ELF *E*-layer noise.

The similarity in the observations made under such different conditions makes us believe that our observations are fairly typical for this phenomenon. In the following we will briefly present what appear to be characteristic features for this *E*-layer noise.

- a: The noise is located in the height interval 90–130 km.
- b: The emissions cover frequencies from about 50 Hz to 3–4 kHz, with maximum amplitude in the lower part of the spectrum. The power spectrum is asymmetrical with a faster roll-off on the low frequency than on the high frequency side (*cf.* Fig. 2.8). This may, however, be somewhat modified by the frequency response of the antenna and electronics. The low frequency limit was, however, also verified by the recordings from the DC channel, which

Fig. 2.8. Power spectrum of ELF E-layer noise recorded by F24. The spectrum is integrated over the total duration of the emissions.



has an AC response up to about 300 Hz (*cf.* Fig. 9.5).

- c: The measured amplitude of the field reaches a maximum of about 5 mV m^{-1} .
- d: There is no marked peak in the amplitude *vs.* frequency representation (Fig. 2.8).
- e: The recorded signal is spin modulated with period equal to half the spin period (Fig. 2.9.) The direction of the electron field is thus well

defined. This is even more clear if the width of the antenna pattern is taken into account. The modulation is most marked at high frequencies, and is more pronounced for the upper 15 km of the region than it is for the fields detected in the lower E-layer. In the last part of the noise-burst, no spin modulation at all can be seen, as if the waves are scattered in all directions and the structure smeared out in the

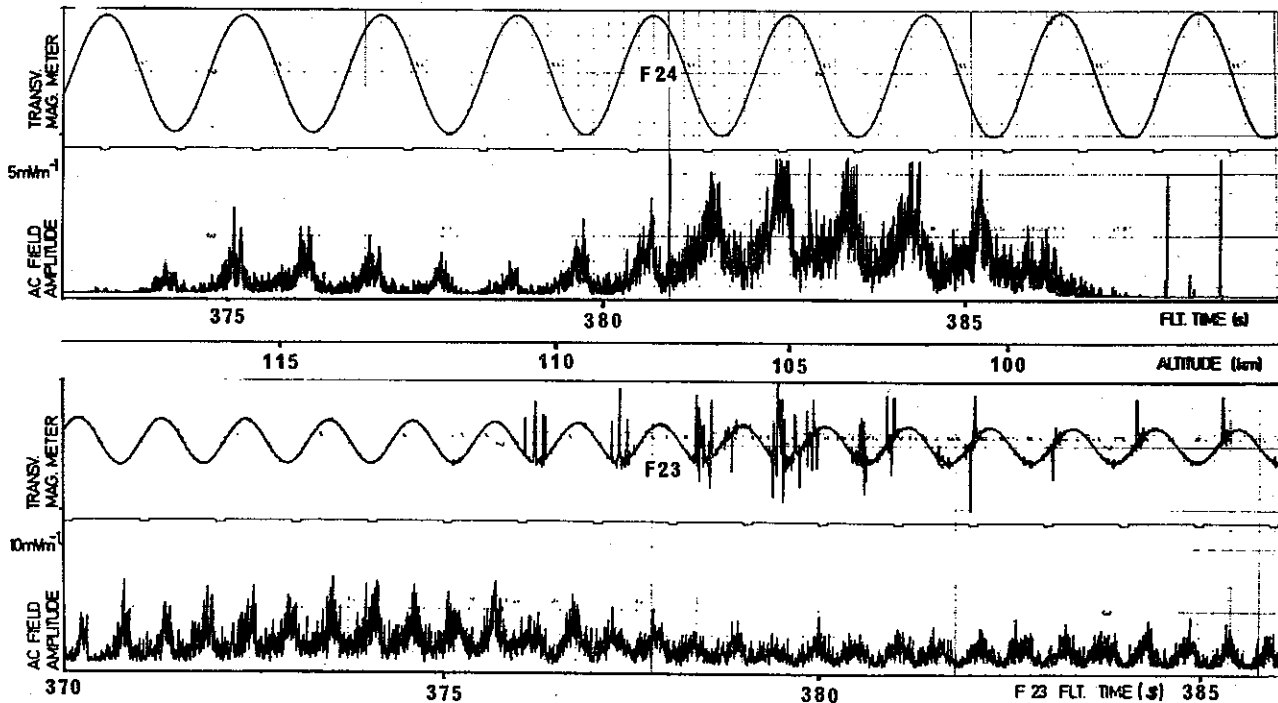


Fig. 2.9. Amplitudes of ELF E-layer noise and signals from the rocket's transverse magnetometer recorded as function of flight time for F24 (upper) and F23 (lower).

denser medium. The thickness of this 'diffusive region' is approximately 4 km (*cf.* Figs. 2.6, 2.7, and 2.9). Much finestructures, with large and rapid fluctuations, are, however, present in the amplitude *vs.* time record (Fig. 2.9). It can also be noted that in some cases the 'heaps' following the rocket spin have a double peak. This means that amplitude maxima occur in two directions close to each other. As this does not appear in all maxima it is supposed to be real effect and not caused by the antenna radiation pattern.

- f: Maximum amplitude and widest frequency range are generally found when the direction of the antenna is near the magnetic east-west direction (*cf.* Fig. 2.10 which is based on data from F24).
- g: The direction of maximum field strength is not constant in space, but can shift with altitude. This can be seen on data from F24 (Fig. 2.10), but is much more pronounced in the F23 measurements (Fig. 2.11) where a shift of approximately 30° occurs. In the F23 measurements difficulties in determining the exact

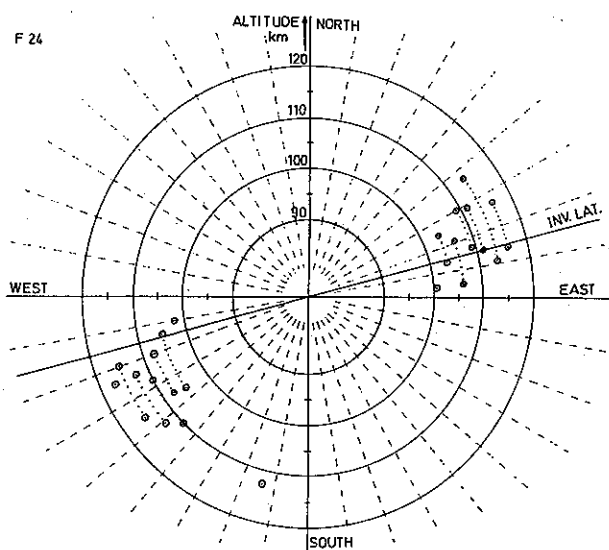


Fig. 2.10. Orientation of the *E*-field antenna in a topographical coordinate system at the times of maximum amplitude in the spin-varying ELF *E*-layer noise bursts (*cf.* Figs. 2.6 and 2.9). Two points connected by a dotted line indicate a double peak in the same burst. Distance from the centre gives the rocket altitude. The line marked INV. LAT. is the direction of constant invariant latitude at 100 km. Data from F24.

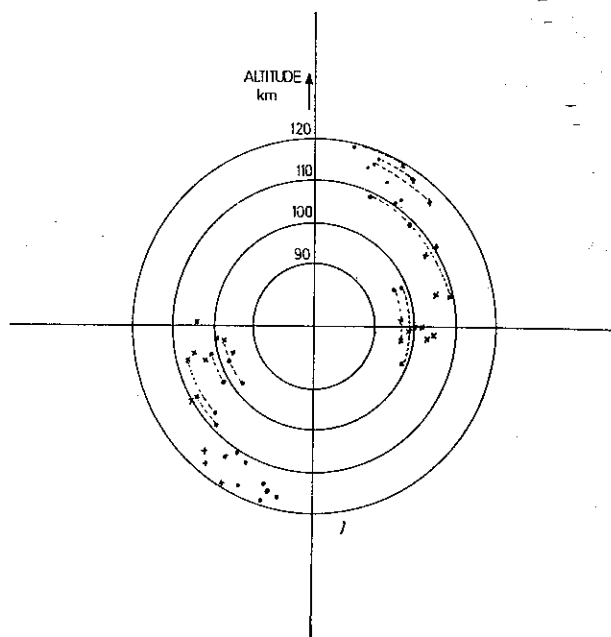


Fig. 2.11. Orientation of the *E*-field antenna in an arbitrary horizontal coordinate system at the times of maximum amplitude in the spin modulation of ELF *E*-layer noise. Points connected by a dotted line indicate multiple peaks in the same burst. Distance from the centre gives the rocket altitude. Data from F23.

orientation in space of the rocket from the magnetic aspect records have made it impossible to locate the electric fields accurately in a topographical coordinate system. However, the changes in direction seen in Fig. 2.11 are quite significant, but the location of the axis is arbitrary. A conspicuous feature in this plot is the division of the points of maximum signal into marked groups, shifted about 30° in direction. The transition takes place quite suddenly around 110 km. It should be noted that at the same altitude the rocket magnetometer showed strong fluctuations, giving evidence of a large local perturbations of the magnetic field (*cf.* Fig. 2.9).

Without listing it as a 'characteristic property' it can be noted that in both cases the enveloping curve for the amplitude *vs.* time (*i.e.* the curve touching the peaks in this record) follows the same pattern and has two marked maxima. This is no artificial effect caused by variations in the rocket attitude. It can also be seen from Figs. 2.6 and 2.7 that the spectral impressions of the emis-

sions in these two bursts are somewhat different. The noise, in the lowest frequencies, is stronger in the second burst, and the spin modulation is less pronounced (*cf.* also point d above). The amplitude minimum of F23 appears approximately where the change in the direction for maximum amplitude takes place.

If we turn to the rocket NASA 14.383, which carried *E*-field instrumentation identical to that in F23 and 24, no *E*-layer noise was observed. The electric field recordings in this rocket were, however, disturbed by some interference, but as this disturbance was not present throughout the flight, sufficient good data were obtained to conclude that this special emission was not present. However, the general disturbance level (*cf.* Table 2.1) was lower at the launch of this rocket than in the two other events, and the rocket particle detectors showed hardly any precipitation of low energy ($E < 15$ keV) particles (Evans, *priv. com.*).

The rocket NASA 18.57 was launched January 16, 1973, 1959 UT, and was fired in the breakup phase of a substorm. Only preliminary data can be presented. However, a very clear case of ELF *E*-layer noise was recorded. The noise seems to be shifted somewhat down in frequency from the earlier cases, but the main characteristics are similar. Preliminary calculated directions of the antenna in the noise maxima are shown in Fig. 2.12. The east-west orientation is also found in this case.

2.5. Other observations of ELF *E*-layer noise

Recently we have become aware of observations made from a rocket, S70/1, flown from Kiruna, February 24, 1970 (Ungstrup, *priv. com.*). The launch took place during an auroral glow situation with weak discrete forms. The rocket was equipped with electric field instrumentation using spherical probes (Danish Space Research Institute). We had the opportunity to study these data; the recordings at *E*-region altitudes appeared to be very similar to those made in F23 and F24. On the upleg, *E*-layer noise was present between 98 and 115 km, and on the downleg from 120 to 99 km. The orientation of the antennas at maximum amplitude was mainly in the east-west direction, in agreement with the observations

from the Andøya rockets. The S70/1 instrumentation also included magnetic antennas. It should be noted that no trace of the *E*-layer noise could be seen in the magnetic records.

Further support for the presence of ELF *E*-layer noise during aurora has been given from two rocket experiments flown in Canada (Maynard, *priv. com.*). In both these cases spin modulated ELF noise was seen when penetrating the *E*-region.

2.6. Discussion of the raw data

As far as the author knows, similar observations have not been reported in the literature. The recordings that bear most resemblance were made by a Japanese group (Iwai *et al.* 1966) and a group at the University of Iowa (Gurnett & Mosier 1969), who observed noise bursts in the same frequency range. It is, however, believed that the noise observed in these cases was not of natural origin, but was generated by an interaction between the rocket and the surrounding plasma (Gurnett & Mosier *loc. cit.*). If we compare these observations with the *E*-layer noise, first of all the altitudes where the noise bursts appear do not agree. While we found this special emission exclusively in the *E*-layer, the other observers detected the noise bursts above 150 and 200 km, respectively, and noticed a systematic decrease in intensity with increasing altitude. This contradicted our findings. Furthermore, the noise showed a very pronounced precession modulation. This effect was not present in our measurements, and the attitudes of the two rockets were also quite different in the periods when the *E*-layer noise was present (*cf.* Figs. 2.3 and 2.5).

The *E*-layer noise thus diverges from the other observations of ELF noise bursts in so many respects that it can be concluded that these two phenomena are of a quite different nature. Electronic payload interference can in our case be excluded as a possible source of the noise. Other possible mechanisms for artificial generation of the field variations (*cf.* Appendix 1) have been examined and rejected.

It can therefore be concluded that the ELF *E*-layer noise is of natural origin. Additional sup-

port for this conclusion is provided by independent observations with different instrumentation, as in the case of our rockets and S70/1 described above.

It is important for the study of the wavefield and its origin to determine whether the waves are electrostatic or electromagnetic. The measurements from S70/1 seem to give evidence of electrostatic waves. This will be further discussed in Chapters 8 and 9.

One should, however, in these measurements be aware of the effects which could be caused by small-scale density gradients in the plasma (*cf.* Appendix I). Such variations will cause fluctuations in the probe contact potentials, which may give rise to detectable potential differences. This would be very difficult to distinguish from real field variations, as the two effects would most probably appear simultaneously. Thus an absolute determination of the amplitude of the field is difficult.

3. CHARACTERISTICS OF THE IONOSPHERIC MEDIUM

3.1. Introduction

For historical reasons the ionosphere is divided into three main parts: The *D*-region situated between approximately 50 and 90 km, the *E*-region continuing upwards to about 150 km, and the *F*-region extending further to about 500 km. Even though there are no sharp divisions between these

layers, it is convenient to make this separation when discussing ionospheric properties. The present work will concentrate on small-scale irregularities in the *E*-region, but to get an impression of what is special for this region and the relations which establish its upper and lower limit, it is necessary to give a brief general view of the ionosphere, up to say 200 km. Our knowledge of ionospheric characteristics was until the mid-1950's derived mainly from ground-based radio and optical measurements. Development of reliable rocket instrumentation during the last two decades has, however, enlarged and revised our knowledge of the ionosphere considerably. New ground techniques, such as radar scattering, partial reflections, and VLF propagation studies, have also contributed with valuable results. However, several uncertainties and question-marks still remain, and the two basic techniques, in situ measurements by rockets and indirect, long-term observations from the ground, which can be 'calibrated' by rocket-measured values, must be joined to give more complete information about the ionosphere.

3.2. The neutral atmosphere and its properties

3.2.1. *Composition of the upper atmosphere.* The composition and distribution of the neutral components in the atmosphere are important not only for the production of the charged species and the ion chemistry, but also for the dynamics of

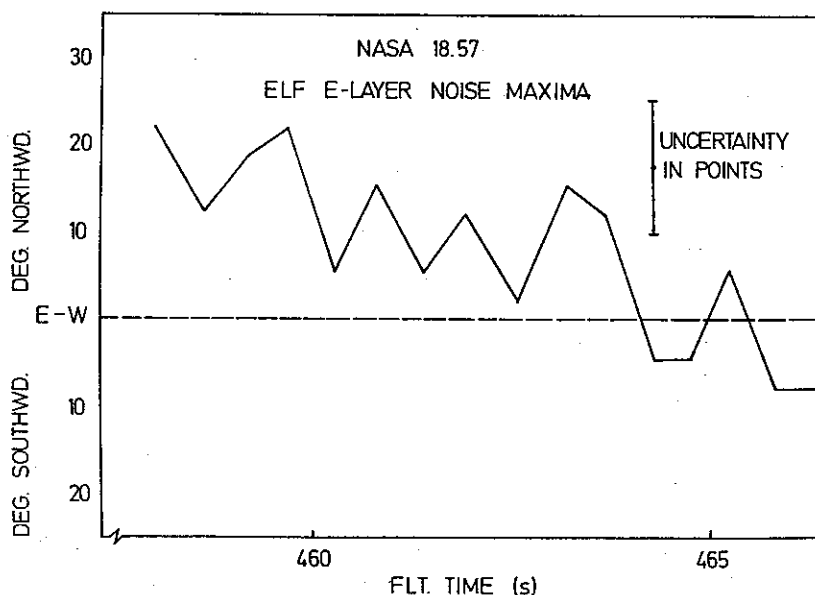


Fig. 2.12. Preliminary plot of direction of antenna in the maxima in the spin-modulated amplitude of ELF *E*-layer noise vs. flight time for the rocket NASA 18.57.

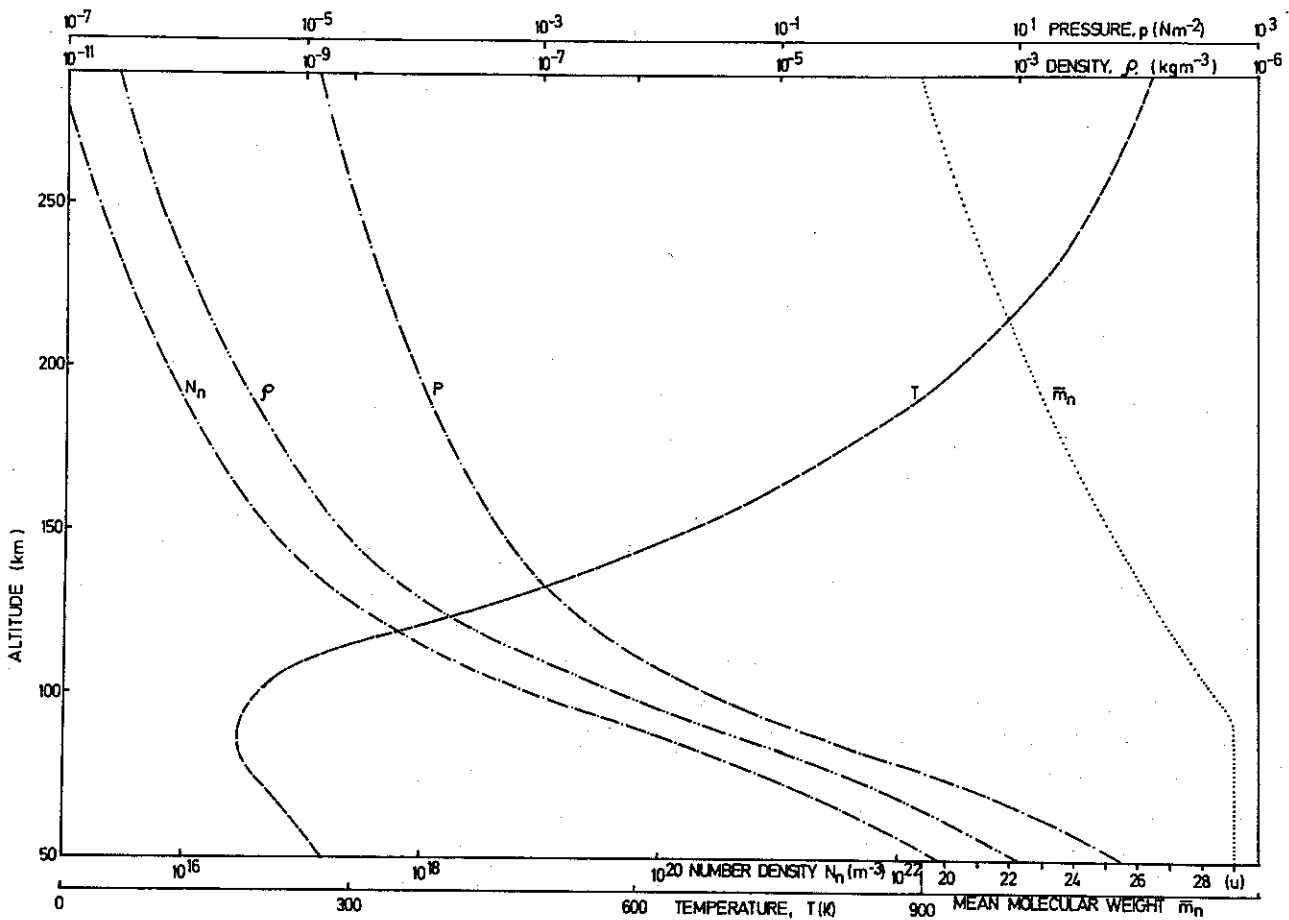


Fig. 3.1. Density (ρ), pressure (p), number density (N_n), mean molecular weight (\bar{m}_n), and kinetic temperature (T_n) for the neutral constituents in a mean atmospheric model (CIRA, 1965).

the gases. Present knowledge about the upper atmosphere is far from satisfactory, especially concerning the minor constituents, some of which are very important for ion chemistry. Since we are more interested in the main features of the ionosphere than in ionic and molecular reactions, this will, however, not disturb our picture considerably. Furthermore, these components will mainly be of importance below 90 km, *i.e.* below the main region of interest in this context.

While direct measurements have given good information on the atmospheric properties up to say 60 km and fairly good up to about 90 km, the experimental data above 90 km are rather sparse. From the available information, standard atmosphere models have been defined; the most commonly used is the COSPAR International Reference Atmosphere Model 1965 (CIRA, 1965), even though this is no longer up to date. Figure 3.1. shows the vertical distribution of pressure, p ,

density, ρ , total number density of neutral constituents, N_n , and mean molecular mass, \bar{m}_n , according to this model.

Up to about 80 km the major constituents are supposed to be present in constant proportions, so that the mean molecular mass will be constant. The dominating properties at these altitudes are: Nitrogen (N_2), which makes up about 78.1%, oxygen (O_2) 21.0%, and argon (A) 0.9%.

While turbulent convection gives a fairly homogeneous gas mixture up to about 80–90 km (the turbopause), there is a rather sudden change at this level, and photochemical dissociation and diffusive separation play the most important roles in disarranging the homosphere distribution above the turbopause.

Figure 3.2. shows the concentration of the total neutral density and the major neutral constituents above 90 km. In the *E*-region N_2 and O_2 are the two dominating molecules, but with increasing

altitude O will be more and more dominating. It should, however, be noted that these concentrations are rather poorly known.

3.2.2. *Transport in the upper atmosphere.* Irregular vertical transport, such as turbulence and eddy motion, is important for mixing the neutral constituents below 80 km. Eddy transport is also important for the H_2O budget of the ionosphere in bringing water vapour up from lower altitudes. Furthermore, turbulence will serve as an 'energy-sink', where larger-scale motions are degraded into small-scale turbulent motions and where the energy is finally lost as heat. Hines (1963) gives as typical parameters for small-scale motions at 80–90 km: size 20 m, velocity 0.5 m s^{-1} , time scale 40 s.

Horizontal transport includes large-scale wind systems. These may be set up by temperature differences, as large zonal wind systems, or driven by pressure gradients, due to daily temperature variations, or be caused by tidal oscillations, global in scale, with periods related to the solar and lunar days.

Smaller-scale effects can be caused by gravity waves with periods of minutes or hours, vertical wavelengths of a few km, and horizontal wavelengths up to thousands of km (Hines 1960). Local transports and irregularities can also be caused by sudden warmings of the lower ionosphere (*e.g.* Belrose 1967). In the *E*-region, large horizontal gradients in the wind velocity, wind shears, are frequently observed, often in connection with sporadic *E*-region ionization (*e.g.* Whitehead 1967).

3.3. Production and loss of ionization

The physical processes controlling the content of charged particles in the ionosphere can be divided into groups labelled production, destruction, and movement of ionization. A continuity equation for the variation of the density of species k , N_k , can be written as

$$\frac{\partial N_k}{\partial t} = q_k - l_k(N_k) - \nabla \cdot (N_k \mathbf{v}_d). \quad (3.1)$$

The symbols q and l represent the production and loss, respectively, and \mathbf{v}_d is the net drift velocity.

For the regions we are considering, *i.e.* below 200 km, the q and l terms dominate, and transport can usually be neglected or treated as a small perturbation. Furthermore, the time constants involved in the processes often make $\frac{\partial N}{\partial t} \ll q_k - l_k(N_k)$, and for slowly varying phenomena we can assume that the equation $q_k = l_k(N_k)$ is adequate.

The principal production process, at least at non-auroral latitudes, for creation of ion-electron pairs is generally accepted to be absorption of solar ultraviolet ($\lambda < 1340 \text{ \AA}$) and X-ray radiation. Although the electron production in the ionosphere involves rather complex processes, a quite good description of many of the main features can be obtained from a simplified theory developed by Chapman (1931). This theory gives a formula for production, q , as a function of height, h , and solar zenith angle, χ , (see *e.g.* Risbeth & Garriott 1969):

$$q(z, \chi) = \frac{\eta I_\infty}{eH} \exp[1 - z - e^{-z} \sec \chi]. \quad (3.2)$$

η is the ionizing efficiency, *i.e.* number of electrons produced per photon absorbed, and z the reduced height $z = \frac{h - h_0}{H}$, h_0 a reference height.

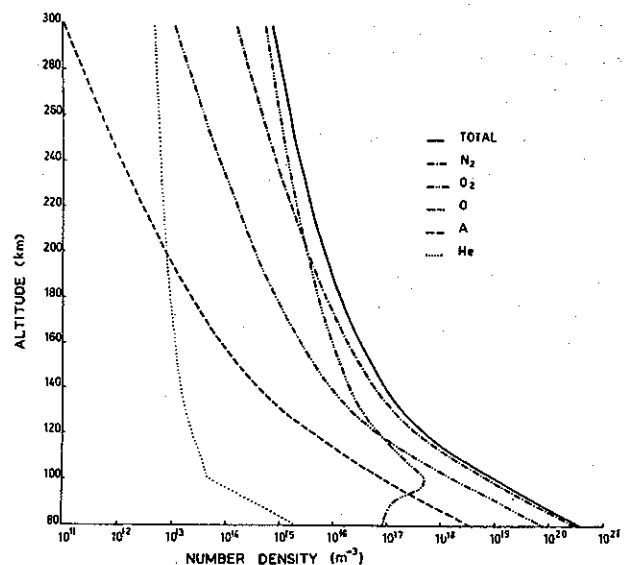


Fig. 3.2. Total density and densities of five major neutral constituents (N_2 , O_2 , O , A and He) as function of altitude for a mean atmospheric model (CIRA, 1965).

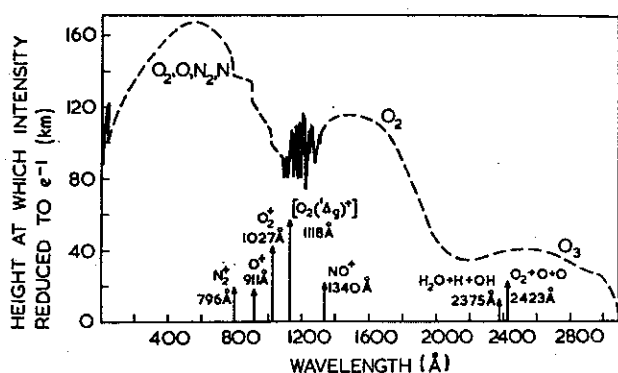


Fig. 3.3. Penetration depth of solar ultraviolet radiation. The curve shows the level where the intensity is reduced to e^{-1} . Wavelengths corresponding to the ionization potential for some important species are marked off (after Thomas 1972).

Several photochemical reactions will tend to neutralize or rearrange the ionization and will thus cause the necessary balance to the production process. The pertinent reactions can be listed as recombination, detachment/attachment, and ion-atom interchange (charge exchange). The most important loss reactions are outlined in Table 3.1. The symbols X,Y,Z, used in the reaction equations may denote either an atom or a molecule. (For more detailed information and references, see *e.g.* Bates 1970.)

To obtain photon-ionization of a species X, symbolized by $h\nu + X \rightarrow X^+ + e$, the photon energy, $h\nu$, must exceed a critical value – the ionization potential of X. This energy threshold also defines a minimum frequency, or maximum wavelength, for the ionizing wave. In Table 3.2, the ionization

Table 3.2. Ionization potentials of the main molecules and atoms in the upper atmosphere and the corresponding upper limit of the wavelength of UV radiation capable of ionizing the species.

Constituent:	O ₂	N ₂	O	N	NO
Ionizing potential, eV	12.08	15.57	13.62	14.54	9.25
Maximum wavelength, Å	1027	796	910	853	1340

potential and corresponding maximum wavelength for the most important ionospheric constituents are given. This, together with the penetration depth in the ionosphere for different wavelengths (Fig. 3.3), will define the processes which can possibly be caused by photon-ionization, if the composition of the neutral atmosphere and the radiation spectrum is known.

Using this we can deduce the following: The production in the lower *D-region* (<70 km) is caused by Lyman- α radiation and cosmic rays. The latter will also give rise to a weak night-time ionization. Between 70 and 80 km ionization of NO by Lyman- α is the most important contribution. In the upper part of the *D-region* (80–90 km), X-rays in the range 1–10 Å will cause ionization of the gases present. In addition, ionization of O₂ in the state (¹ Δ_g) by ultraviolet radiation is suggested as an important process (Henlen & McElroy 1968).

Recombination is believed to be the most important loss process in the *D-region*, at least

Table 3.1. Main reactions for loss of ionization

Recombination:	
Ion-ion	$X^+ + Y^- \rightarrow X + Y$
Electron-ion: Three-body	$X^+ + e + Z \rightarrow X + Z$
Radiative	$X^+ + e \rightarrow X + h\nu$
Dissociative	$XY^+ + e \rightarrow X + Y$
Detachment (\rightarrow), attachment (\leftarrow):	
Collisional d., three body a.	$X^- + Z \rightleftharpoons X + e + Z$
Associative d., dissociative a.	$X^- + Y \rightleftharpoons XY + e$
Photo d., radiative a.	$X^- + h\nu \rightleftharpoons X + e$
Ion-atom interchange:	
	$X^+ + YZ \rightarrow XY^+ + Z$
	$X^+ + YZ \rightarrow YZ^+ + X$

during the day, but due to the presence of negative ions, the loss reactions at these altitudes are very complicated.

The predominant ions are NO^+ , O_2^+ , and, in the upper part, N_2^+ . Water vapour complexes and negative ions are important for the chemistry in the region, but little is known about their concentrations.

The variations in the undisturbed daytime *E*-region can be described by a Chapman production function and with recombination as the dominating loss process. (This combination is often called an α -layer.) The ion production is mainly caused by photon-ionization from X-rays, 10–100 Å, and UV radiation in the 1000–1500 Å region. The chief ions are believed to be O_2^+ , NO^+ , N_2^+ , and O^+ . Recent rocket measurements have also shown that metallic ions may exist in appreciable quantities, often in very narrow layers (Chen & Harris 1971).

Although the general behaviour of the layer may be described by the α -layer function, a number of disparities have been observed. One of these is the existence of a night-time *E*-layer where the electron density does not fall below 10^9 m^{-3} (Chen & Harris 1971), while, according

to the Chapman theory, the layer should vanish during night-time. This means that either the effective *E*-region recombination rate must change at night (Yonezawa 1968), or that a source of ionization exists. As a possible source for the night-time *E*-region, Lyman- β scattered to the dark side of the earth has been suggested (Swider 1965; Ogawa & Tohmatsu 1966). Transport effects, which to a great extent have been neglected in *E*-region studies, may also be responsible for several of the irregular features. However, the most important ionization source during disturbed conditions at high latitudes is found to be precipitating particles (see Sect. 3.4).

The ionization of the lower part of the *F*-region (150–200 km) is mainly caused by radiation in the range 170–910 Å, and the behaviour of the region is close to that of an α -layer. The primary ions are O^+ and N_2^+ .

Table 3.3 and Figs. 3.4 and 3.5 give a summary of this section. The table lists the various production and loss processes and their importance in the different regions, while the figure gives the results of these reactions, *i.e.* the densities of the most important charged species in the lower ionosphere.

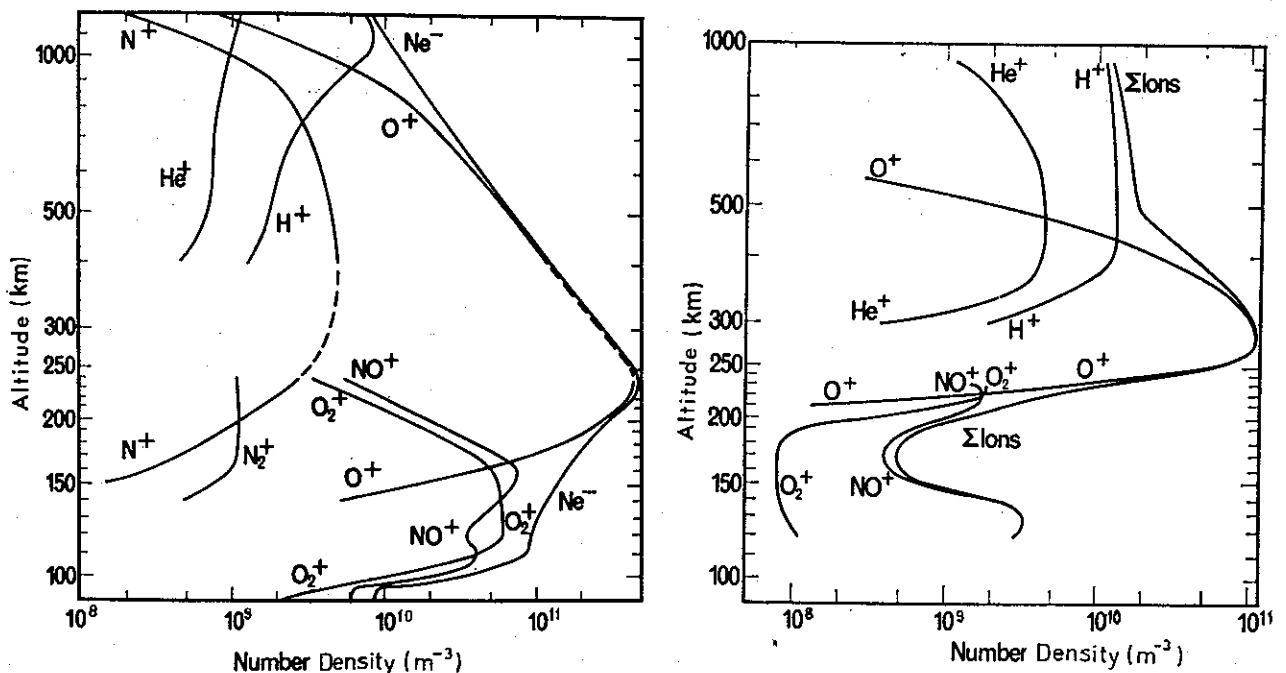


Fig. 3.4. Ionic composition of IQSY daytime (left) and night-time (right) ionosphere above 100 km (after Johnson 1969).

Table 3.3. Production and loss functions in the ionosphere

Process	Contribution to Eq. 3.1.	D-region Approx. 50-90 km	E-region Approx. 90-150 km	FI-region Approx. 150-200 km
PRODUCTION				
Solar photon-ionization	$q(h)$	Ly- α (1216 Å) \rightarrow NO X-rays (1-10 Å) \rightarrow N ₂ , O ₂ UV(1027-1118 Å) \rightarrow O ₂ (¹ / ₂ g)	Ly cont. (800-910 Å) \rightarrow N ₂ , O ₂ Ly- β (1026 Å) \rightarrow O ₂ UV($\lambda < 1027$ Å) \rightarrow O ₂ , N ₂ , O X-rays (10-170 Å)	EUV (170-911 Å) $\lambda < 910$ Å \rightarrow O $\lambda < 796$ \rightarrow N ₂
Ionization by particles (Mostly at high latitudes)	$q(h)$	Electrons > 30 keV Protons > 2MeV (PCA) Cosmic rays	Electrons 1-30 keV Very important	Electrons $\lesssim 1$ keV (most important at night)
LOSS				
Ion-ion recombination	$\alpha N_+ N_-$	Important	Few. neg ions	Neg. ions can be neglected
Electron-ion recomb. Three-body	$\alpha_e N_+ N_e$	Important	Insignificant	Not important
Radiative	$\alpha_e = \alpha_e(h)$	Insignificant	Not important	Not important
Dissociative		Important	Principal loss mech.	Principal loss mech.
Attachment Radiative	$\beta_{eff} N_e$	Important	Can maintain some neg. ions at night	Weakly involved
Three body		Most important		
Dissociative				
Detachment Collisional	$\delta(h) N_e, \kappa N_-$	Important	Weakly involved	Insignificant
Associative		Important, causes day/night change of N_-/N_e	Effective by day	Causes absence of neg. ion
Photo		Weakly involved	Important	Important
Ion-atom interchange	$\gamma(h) N_+$			

3.4. Ionization by precipitating particles

Energetic particles precipitating into the atmosphere collide with neutral atoms and molecules and gradually lose their energy to these. A part of this energy goes to ionization – and excitation – of the neutral air. The approximate penetration depths for protons and electrons as a function of energy are shown in Fig. 3.6. As the collision process, and thus the penetration, is a statistical process, the values in this figure should be considered as average heights where most of the energy of the primary particles is lost. It can be seen that only electrons and protons with initial energies greater than 20 and 200 keV, respectively, can cause ionization at *D*-region altitudes. The *E*-region will mainly be influenced by 0.5–10 keV electrons and 10–200 keV protons, while the *F*-region ionization mainly comes from electrons with energy less than 1 keV. Figure 3.7 shows the energy loss rate for electrons with initial energy 6 keV as a function of altitude. It is seen that these particles lose almost all their energy within the *E*-region.

At low and medium latitudes only the *D*-region seems to be affected by ionization from precipitating particles (Whalen *et al.* 1971). This is also in accordance with the statistical picture of the particle spectrum *vs.* latitude (*e.g.* Søråas 1973).

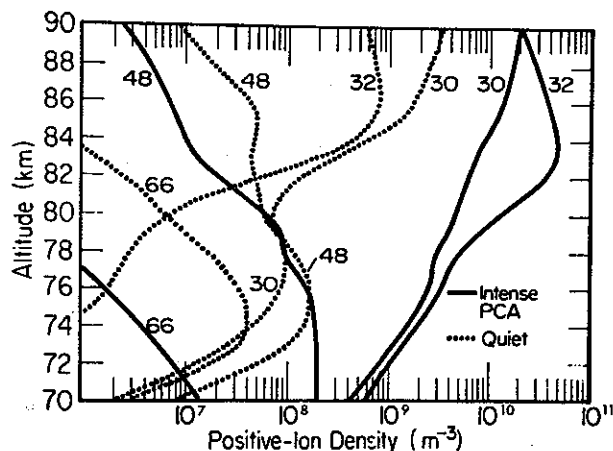


Fig. 3.5.a. Altitude profiles for molecular positive ions and hydrates during quiet and intense PCA conditions (after Reid 1972). Mass identifications are: 30 = NO^+ , 32 = O_2^+ , 48 = $\text{NO}^+ \cdot \text{H}_2\text{O}$, 66 = $\text{NO}^+ \cdot 2(\text{H}_2\text{O})$.

While particle ionization at these latitudes plays a minor role, in the auroral zone during disturbed conditions it may be the main source of ionization. The energy spectrum of the precipitating particles on the night side of the auroral oval implies that the ionization will normally be most enhanced in the *E*-region and upper *D*-region; *i.e.* from approximately 80 to 150 km (Hultqvist 1972). Here the peak electron density may typically increase from say 10^9 – 10^{10} m^{-3} during quiet conditions to 10^{12} m^{-3} during an auroral event (McNamara 1969).

On the dayside, precipitating particles will mainly cause ionization in the *D*- and *E*-region equatorwards of the auroral oval, while the particle spectrum on the dayside of the oval, the polar cusp, is considerably softer, with a correspondingly increased altitude for the ionization. Inside the polar cap the particle spectrum is basically soft, and during normal conditions only the upper ionosphere will be affected by precipitation ionization.

This pattern will be changed by irregular and sporadic precipitation and during irregular precipitation events with great fluxes of high energy particles (*e.g.* Polar Cap Absorption (PCA) events) when significant ionization can be present in the lower *D*-region from the polar cap to $\lambda \sim 50^\circ$.

To conclude it can be said that ionization by

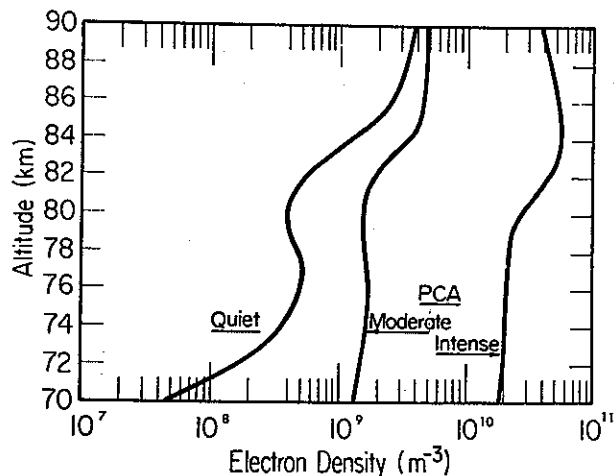


Fig. 3.5.b. Electron concentrations below 90 km during quiet conditions and during moderate and intense PCA conditions (after Reid 1972).

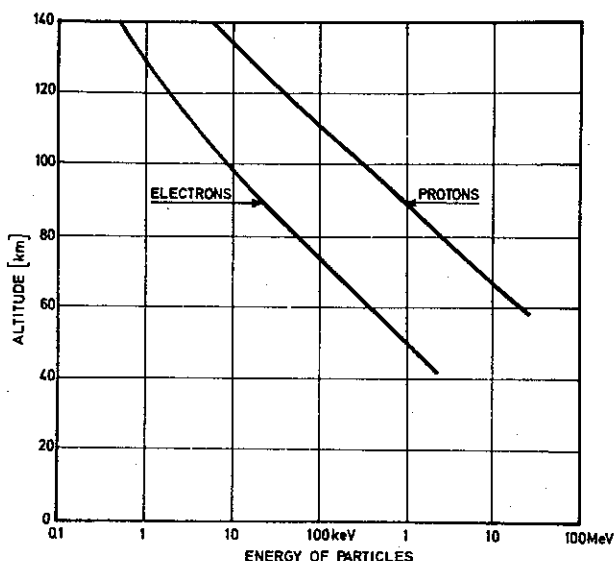


Fig. 3.6. Penetration depths in the earth's atmosphere as function of particle energy for vertically incident electrons and protons (after Mæhlum 1973).

particles plays a very important role in the production of the high latitude ionosphere, especially in the auroral E-region. Because of the great variations, both in time and space, which are present in the particle fluxes, this will involve large fluctuations in both the electron density and its altitude distribution.

3.5. Ionospheric parameters

In addition to composition and density several other parameters which can be measured directly

or otherwise calculated are important in describing the behaviour of the plasma.

3.5.1. Temperatures of the ionospheric constituents.

In dealing with ionospheric plasma it is necessary to distinguish between the temperature of the neutrals, T_n , and electron and ion temperatures, T_e and T_i . Assuming a Maxwellian velocity distribution, temperature is usually defined in terms of the mean energy, $E = \frac{3}{2}\kappa T$, where $\kappa =$ Boltzmann's constant.

It is widely accepted that up to 120–130 km thermal equilibrium is established, *i.e.* $T_e = T_i = T_n$, during undisturbed conditions. The numerical values of the temperature are, however, subjected to daily and latitudinal variations. At greater altitudes selective heating (Bowhill & Geisler 1966) and reduced collision rates lead to temperature differences, so that $T_e > T_i > T_n$. As energy transfer through collisions is most effective among particles of equal masses, the T_i and T_n curves will separate at higher altitudes than those for T_e and T_n . Figure 3.8 shows the altitude variation of T_n after CIRA (1965) and reasonably representative curves for T_i and T_e according to Evans & Mantas (1968) for quiet conditions. The picture, with $T_e \approx T_i$ in the D- and lower E-region, will be considerably changed, especially at high latitudes, during disturbed conditions (*cf.* the T_a -curves in Fig. 3.8). Precipitation of energetic particles will not only cause in-

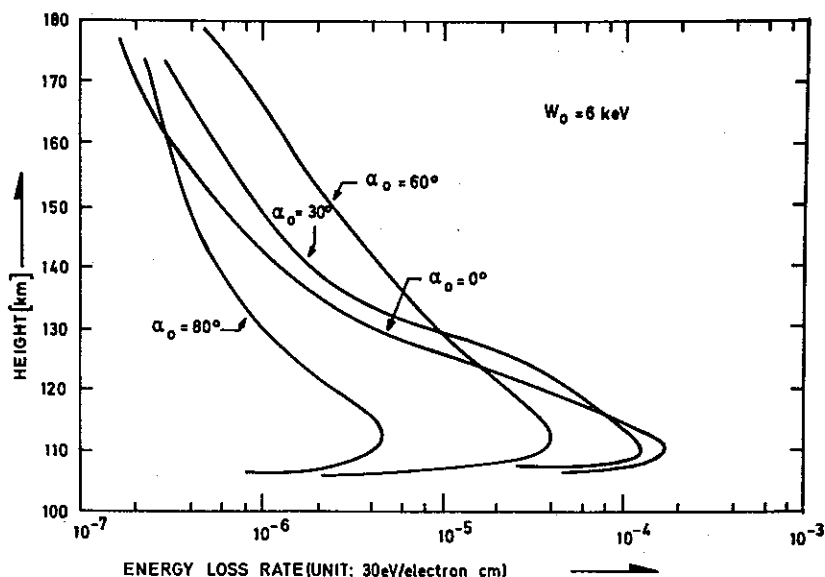


Fig. 3.7. Energy loss rate profiles for electrons with initial energy 6 keV and with impact angle α (after Wedde 1970).

creased temperatures at the penetration altitudes, but also change the ratio T_e/T_i , and lower the altitude where this ratio becomes larger than one (cf. e.g. McNamara 1969).

3.5.2. *Collision frequencies.* Collision processes are very important in the ionospheric plasma both for energy transfer and because they will have influence on and modify several of the other plasma parameters. In the study of ionospheric irregularities, collisions are involved both in the generation and motion of the irregularities, and it is therefore necessary to have information about variations in the collision rates for the various species. A thorough discussion of collision phenomena in atmospheric gases is very complicated and beyond the scope of this paper. However, for most applications simplified methods can be used.

Under the term 'collision' several processes involving energy transfer between particles, such as elastic ('classical') collisions, scattering in a Coulomb field, ionization, dissociation, excitation of particular states, etc., can be involved. For a constituent, k_1 , each interaction process, γ , has a specific cross section, $\sigma_{k_1k_2}(\gamma, v)$, which is also a function of velocity, v , and target particle, k_2 , with its corresponding collision frequency, $\nu_{k_1k_2}$. In ionospheric studies it is convenient to introduce

a weighted cross section, $\sigma_{k_1k_2}^m$, where change in momentum transfer with scattering angle is taken into account, and a corresponding 'momentum transfer collision frequency', $\nu_{k_1k_2}^m$. As the cross sections depend on the velocity, 'effective' collision frequencies, $\langle \nu_{k_1k_2}^m \rangle$, are obtained by averaging over the velocity distribution.

The collision processes which will predominate in this context are:

- i) Elastic collisions between electrons and neutrals, $\langle \nu_{en}^m \rangle$.
- ii) Elastic collisions between ions and neutrals, $\langle \nu_{in}^m \rangle$.
- iii) Coulomb collisions between electrons and ions, $\langle \nu_{ei}^m \rangle$.

From a general expression of the rate of energy transfer between two gases with Maxwellian velocity distributions (Desloge 1962), Banks (1966 *a, b*) derived the following expression for the average momentum transfer collision frequency between species k_1 and k_2

$$\langle \nu_{k_1k_2}^m \rangle = \frac{4}{3} N_{k_2} \left(\frac{8\kappa}{\pi} \right)^{1/2} \left[\frac{T_1}{m_1} + \frac{T_2}{m_2} \right]^{1/2} \bar{Q}_m, \quad (3.3)$$

where \bar{Q}_m , the average momentum transfer cross section for the reaction between k_1 and k_2 is defined as:

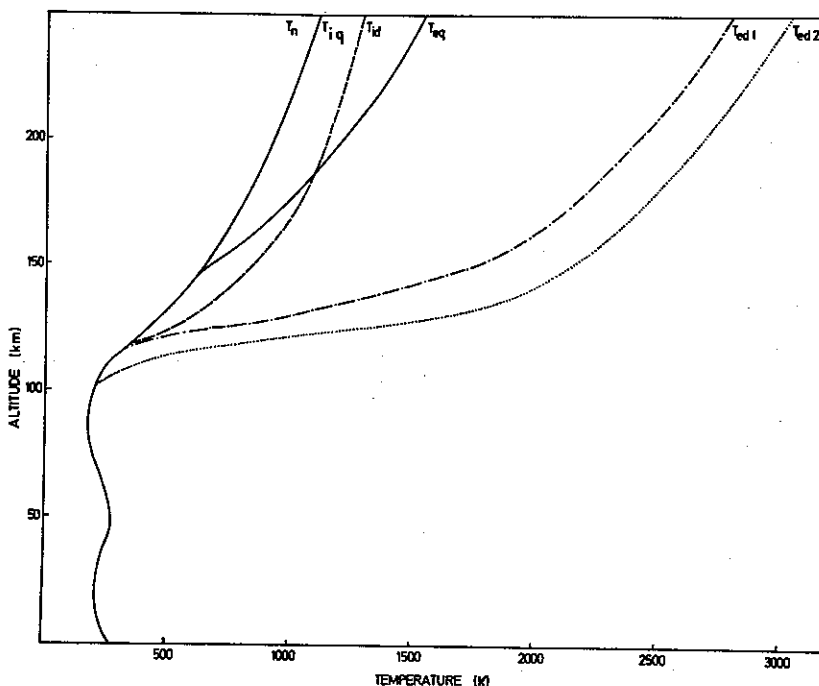
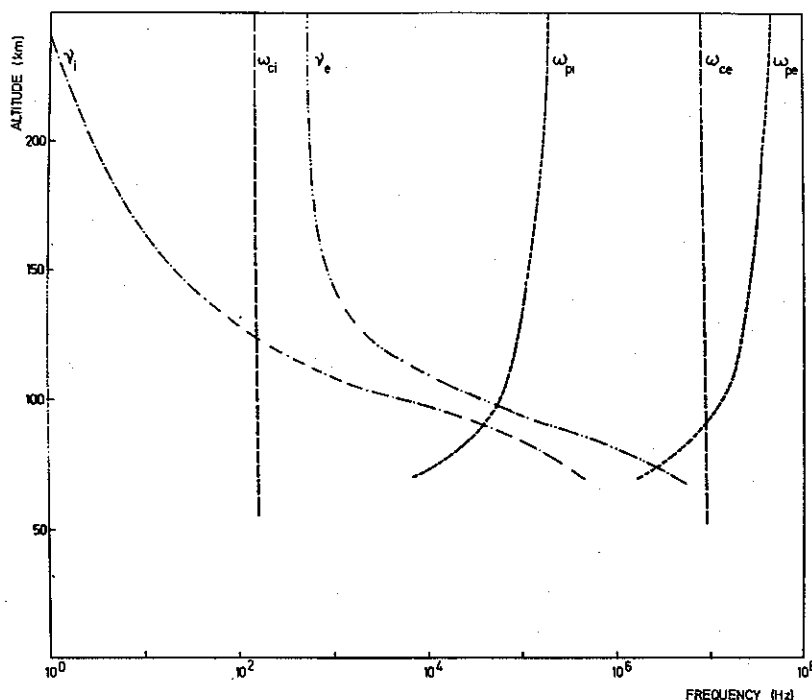


Fig. 3.8. Height profiles for ionospheric temperatures. T_n is the temperature of neutral constituents (after CIRA 1965), T_{eq} and T_{iq} the electron and ion temperatures for quiet conditions (after Evans & Mantas 1968). T_{ed1} and T_{id1} are electron and ion temperatures for moderately disturbed conditions and T_{ed2} electron temperatures for strong disturbance (after Walker & Rees 1968).

Fig. 3.9. Variations with altitude of electron and ion collision frequency (ν_e and ν_i), electron and ion gyro frequency (ω_{ce} and ω_{ci}) and electron and ion plasma frequency (ω_{pe} and ω_{pi}) calculated for a mean model ionosphere.



$$\bar{Q}_m = K^3 \int_0^{\infty} v^5 \sigma_{k_1 k_2}^m \exp(-Kv^2) dv, \quad (3.4)$$

with

$$K = \left[\frac{2\kappa T_1}{m_1} + \frac{2\kappa T_2}{m_2} \right]^{-1}.$$

For the lower ionosphere, where our main interest is directed, rough numerical values for the collision frequencies can be obtained from the simplified expressions (Boström 1972)

$$\{\nu_{en}\} = 1.5 \cdot 10^{-17} \{N_n\} \{T\}, \quad (3.6)$$

$$\{\nu_{ei}\} = [59.0 + 1.82 \ln(\{T^3\}/\{N_e\})] \cdot 10^{-6} \{N_e\} \{T^{-3/2}\}, \quad (3.7)$$

$$\{\nu_{in}\} = 4.2 \cdot 10^{-16} \{N_n\}. \quad (3.8)$$

These are only numerical equations, and symbols in $\{\}$ represent the values of the various parameters measured in the SI system.

The relative importance of the two different collision processes for electrons will vary with altitude according to variations in composition and temperatures. In an attempt to appreciate the relative importance of electron-neutral and electron-ion collisions, Banks (1966a) found that the

latter will make a significant contribution to the total collision frequency when $\frac{N_+}{N_n} > \frac{T_e^2}{1.3 \cdot 10^{11}}$.

From the values of densities and temperatures given earlier in this chapter it is found that electron-ion collisions will be important in the upper part of the *E*-region and above this region.

Collision frequencies calculated from Eqs. (3.6), (3.7), and (3.8) using ionospheric parameters given earlier in this chapter, are shown in Fig. 3.9.

3.5.3. Characteristic frequencies and dimensions in the ionospheric plasma. The introduction of a magnetic field, B , will influence the plasma in several ways. The first detectable effect is that it will cause a gyrating motion of the charged particles, the frequency of which is given by

$$\omega_{ck} = \frac{Z_k e B}{m_k}. \quad (3.9)$$

The ratio ω_{ck}/ν_k indicates whether the particle motion is dominated by the magnetic field or if collision processes play the most important role. Figure 3.9 gives the altitude variation of the gyro frequency for electrons and an ion with mass 30 together with the collision frequencies for the same species as a function of altitude. The

magnetic field for an auroral station (Tromsø) has been used in the ω_{ck} calculations.

Another characteristic frequency is the so-called plasma frequency, defined by

$$\omega_{pk} = \left(\frac{N_k Z_k^2 e^2}{\epsilon_0 m_k} \right)^{1/2} \quad (3.10)$$

The magnitude of the plasma frequency characterizes the rate at which the electrostatic restoring forces in a plasma eliminate deviations from neutrality. The height variation of the ionospheric electron and 'mass-30-ion' plasma frequency is also shown in Fig. 3.9.

Fig. 3.10 gives the magnitudes of several length parameters pertaining to the ionospheric plasma. They will be defined in the following.

The radius of the gyrating motion of a k -particle, the cyclotron radius (or Larmor radius) r_{ck} , is

$$r_{ck} = \frac{v_{\perp}}{\omega_{ck}} = \frac{v_{\perp} m_k}{Z_k e B} \quad (3.11)$$

where v_{\perp} is the component of the particle velocity transverse to the magnetic field. On average, if the velocity distribution is isotropic, v_{\perp} can be set equal to the mean thermal velocity for Maxwellian distribution, and the following expression for the mean cyclotron radius is obtained

$$\langle r_{ck} \rangle = \left(\frac{8 m_k \kappa T_k}{\pi Z_k^2 e^2 B^2} \right)^{1/2} \quad (3.12)$$

A k -particle moving with velocity v_k will on the average move a distance $\frac{v_k}{\nu_k}$ between each collision.

This distance, the mean free path, λ_{vk} , will for a Maxwellian velocity distribution not correspond exactly to the expression obtained when putting $v_k = \langle v_k \rangle$. One will instead get (Morse 1964)

$$\lambda_{vk} = \left(\frac{\pi \kappa T_k}{2 m_k \nu_k^2} \right)^{1/2} \quad (3.13)$$

As for the ratio ω_{ck}/ν_k , the ratio r_{ck}/λ_{vk} , will contain information about the influence of the magnetic field relative to collisions on the particle motion. If, for instance, $r_{ck}/\lambda_{vk} \ll 1$, trajectories of the k -particles can be regarded as almost straight lines between consecutive collisions without being significantly affected by the magnetic

field, and the diffusion in the plasma can be considered isotropic.

The length parameter corresponding to the plasma frequency is the Debye length, λ_D , which for a non-isothermal plasma is defined as (see Frank-Kamenetski 1967)

$$\lambda_D = \left[\frac{e^2}{\epsilon_0 \kappa} \left(\frac{\sum_i Z_i^2 N_i}{T_i} + \frac{N_e}{T_e} \right) \right]^{-1/2} \quad (3.14)$$

Subscripts i refer to ions of type i , and a summation is made over all ionic species.

In application, the effects from the ions are usually neglected, and this gives the following simple expression for the Debye length:

$$\lambda_D = \left(\frac{\epsilon_0 \kappa T_e}{e^2 N_e} \right)^{1/2} = \left(\frac{\kappa T_e}{m_e} \right)^{1/2} \frac{1}{\omega_{pe}} \quad (3.15)$$

The Debye length is thus the distance a particle with velocity $\left(\frac{\kappa T_e}{m_e} \right)^{1/2}$ will travel within a time interval $\frac{1}{\omega_{pe}}$. This velocity is frequently used as the electron thermal velocity instead of the mean velocity in the Maxwell distribution, $\left(\frac{8}{\pi} \frac{\kappa T_e}{m_e} \right)^{1/2}$.

The Debye length is also often called the 'Debye

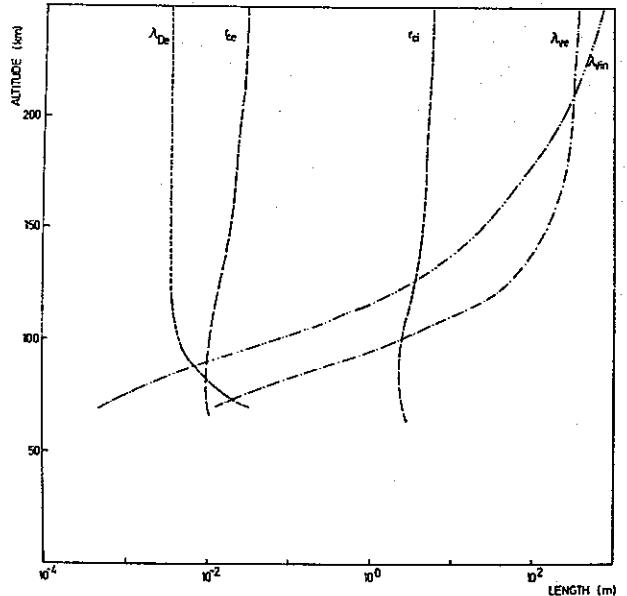


Fig. 3.10. Altitude profiles of electron and ion gyro-radius (r_{ce} and r_{ci}), electron and ion mean free path (λ_{ve} and λ_{vi}), and electron Debye length (λ_{De}) for a mean model ionosphere.

shielding distance'. The charged particles in a plasma have a tendency to rearrange themselves in such a way as to shield any electrostatic fields due either to a surface at some non-zero potential or to a charge within the plasma. In the 'classical' approach it appears that this rearrangement of charged particles effectively cancels out any electrostatic fields within a distance of the order of λ_D . However, in several cases, for instance in a plasma far from thermal equilibrium, the Debye length is not always well defined, and it has also been shown (Öpik 1965) that the effective shielding in practical circumstances may deviate considerably from the classical Debye length.

4. ELECTRODYNAMICS OF THE IONOSPHERE

4.1. Introduction

The introduction of an electric field in the ionosphere will set charged particles into motion, and cause a drift of the ionization as a whole as well as of ionization irregularities, and electric currents will be flowing. However, the relation between the direction and magnitude of the field and the resulting particle velocity is not a simple one, as effects caused by the magnetic field and collisions will give different mobilities in different directions. In this chapter we will not discuss the origin of the electric field, but just postulate its presence and see how the motion of the charged species is influenced.

Instead of starting out with the equations of motion and calculate the mobility of the particles, we will follow a more qualitative, physical derivation, making use of a procedure given by *e.g.* Baker & Martyn (1953), Ratcliffe (1959*a*), and Boström (1973). This method has the advantage that it more easily gives an impression of how variations of the different parameters result in changes in directions and magnitudes of velocities and currents.

4.2. Particle motions and conductivities

In the presence of a magnetic field an electric field is defined only with reference to a coordinate system. From one frame of reference to another both \mathbf{E} and \mathbf{B} will change according to Lorentz

transformations (*cf. e.g.* Møller 1952). If we operate in a coordinate system where the electric and magnetic fields are \mathbf{E} and \mathbf{B} , an object with velocity \mathbf{v} will experience a field \mathbf{E}' , given by

$$\mathbf{E}' = \mathbf{E} + \mathbf{v} \times \mathbf{B}, \quad (4.1)$$

in non-relativistic formulation. This field will cause a drift of the particles which in turn will give rise to an electric current. In analogy to the simple Ohm's law for the current density in a linear conductor, one can introduce a 'generalized Ohm's law'

$$\mathbf{j} = \underline{\underline{\sigma}}(\mathbf{E} + \mathbf{v} \times \mathbf{B}). \quad (4.2)$$

Here the conductivity, $\underline{\underline{\sigma}}$, is a tensor, where the tensor elements depend on the mobility of the particles.

Between collisions the motion of a k -particle will only be affected by the electric and magnetic field, and its equation of motion will be

$$m_k \frac{d\mathbf{v}_k}{dt} = q_k(\mathbf{E} + \mathbf{v}_k \times \mathbf{B}). \quad (4.3)$$

If the collision frequency is ν_k the probability of a collision taking place in the time interval from t to $t + dt$ will be $\nu_k \exp[-\nu_k t] dt$. From the equation of motion the velocities in the collisionless case can be found, and from this also the mean velocity components when collisions are taken into account (*cf. e.g.* Boström 1973). In the case where \mathbf{B} is directed along the negative z -axis and \mathbf{E} is in the xz -plane these are

$$v_{kx} = \frac{\varepsilon_k \nu_k \omega_{ck}}{\nu_k^2 + \omega_{ck}^2} \frac{E_x}{B}, \quad (4.4a)$$

$$v_{ky} = \frac{\omega_{ck}^2}{\nu_k^2 + \omega_{ck}^2} \frac{E_x}{B}, \quad (4.4.b)$$

$$v_{kz} = \frac{q_k}{m_k \nu_k} E_z. \quad (4.4.c)$$

Taking the motion of the neutral gas into account, Eqs. (4.4) can be presented in the more general form

$$\begin{aligned} \mathbf{v}_k - \mathbf{v}_n = & \frac{\varepsilon_k \nu_k \omega_{ck}}{\nu_k^2 + \omega_{ck}^2} \frac{\mathbf{E}_\perp + \mathbf{v}_n \times \mathbf{B}}{B} \\ & + \frac{\omega_{ck}^2}{\nu_k^2 + \omega_{ck}^2} \frac{(\mathbf{E}_\perp + \mathbf{v}_n \times \mathbf{B}) \times \mathbf{B}}{B^2} + \frac{q_k}{m_k \nu_k} \mathbf{E}_\parallel. \end{aligned} \quad (4.5)$$

Subscripts \perp and \parallel refer to components perpendicular and parallel to the magnetic field, and $\epsilon_k \equiv q_k/|q_k|$.

The difference between the electron and ion velocities will result in an electric current. Assuming charge neutrality and neglecting effects from negative ions, *i.e.* setting $N_e = N_i$, the current density becomes

$$\mathbf{j} = N_e e(\mathbf{v}_i - \mathbf{v}_e). \quad (4.6)$$

Combination of Eqs. (4.2) and (4.5) gives the following expression for the current density:

$$\mathbf{j} = \sigma_P(\mathbf{E}_\perp + \mathbf{v}_n \times \mathbf{B}) + \sigma_H \mathbf{B} \times (\mathbf{E}_\perp + \mathbf{v}_n \times \mathbf{B})/B + \sigma_\parallel \mathbf{E}_\parallel. \quad (4.7)$$

The conductivity tensor has thus the form

$$\tilde{\sigma} = \begin{pmatrix} \sigma_P & \sigma_H & 0 \\ -\sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_\parallel \end{pmatrix} \quad (4.8)$$

The three different tensor elements, σ_P , σ_H and σ_\parallel , called the Pedersen, Hall, and parallel conductivity, respectively, are found to be

$$\sigma_P = \left(\frac{v_e \omega_{ce}}{v_e^2 + \omega_{ce}^2} + \frac{v_i \omega_{ci}}{v_i^2 + \omega_{ci}^2} \right) \frac{eN_e}{B}, \quad (4.9.a)$$

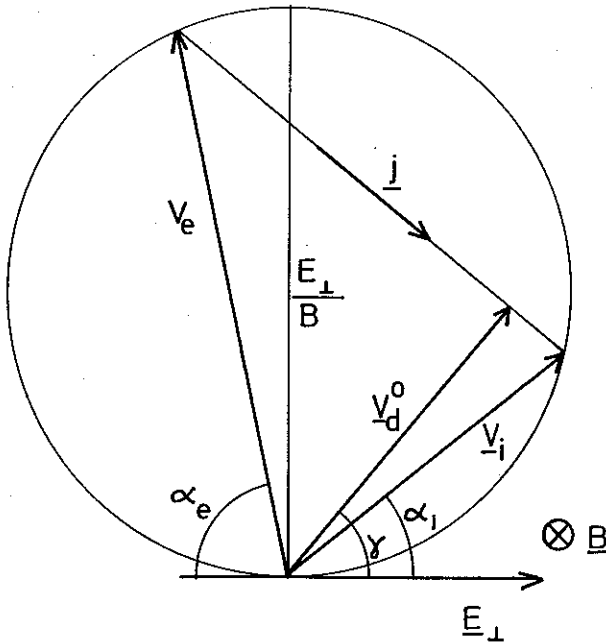


Fig. 4.1. Graphical construction of direction and magnitudes of drift velocities of electrons (\mathbf{v}_e) and ions (\mathbf{v}_i) due to an electric field E_\perp . The resulting current is in the direction of \mathbf{j} , neutral ionization drift in the \mathbf{v}_d^0 direction.

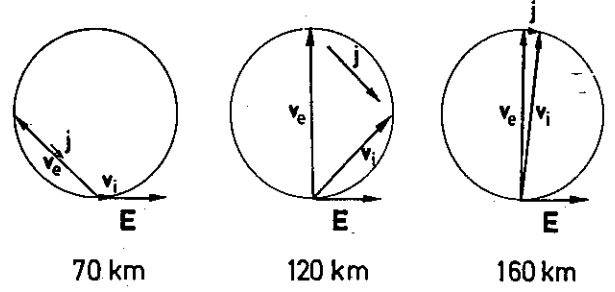


Fig. 4.2. The direction of the drift velocities and electric current, \mathbf{j} , relative to the electric field, \mathbf{E} , at three different heights of the ionosphere (Bostrom 1973).

$$\sigma_H = \left(\frac{\omega_{ce}^2}{v_e^2 + \omega_{ce}^2} - \frac{\omega_{ci}^2}{v_i^2 + \omega_{ci}^2} \right) \frac{eN_e}{B}, \quad (4.9.b)$$

$$\sigma_\parallel = \left(\frac{1}{m_e v_e} + \frac{1}{m_i v_i} \right) e^2 N_e. \quad (4.9.c)$$

The perpendicular conductivities, σ_P and σ_H , appear as a result of the collisions and will be related to currents parallel (σ_P) and perpendicular (σ_H) to \mathbf{E}_\perp . On the other hand, collisions will reduce σ_\parallel .

The directions and magnitudes of the transverse velocities and current may be visualized as in Fig. 4.1. The magnetic field is directed into the paper plane, and the x -axis is in the direction of \mathbf{E}_\perp . The velocity vectors will then form the angles $\alpha_k = \arctan \frac{v_{ky}}{v_{kx}}$ with the x -axis. From Eqs. (4.4) these are found to be

$$\alpha_k = \arctan \frac{\omega_{ck}}{v_k}. \quad (4.10)$$

The diameter of the circle circumscribing the triangle formed by the vectors \mathbf{v}_e and \mathbf{v}_i is

$$D = \frac{v_k}{\sin \alpha_k} = \frac{v_k^2}{v_{ky}} = \frac{E_\perp}{B}, \quad (4.11)$$

and the velocities can thus also be expressed as

$$v_k = \frac{E_\perp}{B} \sin \alpha_k. \quad (4.12)$$

The current will be in the direction of the line between the arrowheads of \mathbf{v}_e and \mathbf{v}_i .

In this way we can, by making a circle with diameter $\frac{E_\perp}{B}$ and drawing two chords in the directions α_e and α_i , get an impression of the

motion of the charged particles and the direction and magnitude of the current density.

Figure 4.2 shows these properties for three different heights. It can be seen how the directions, with a constant electric field, will change, corresponding to a change in the ratio between the Pedersen and Hall conductivities. In Fig. 4.3 the variation of the conductivities for a normal nighttime ionosphere, based on parameters given in Chapter 3, is plotted. From these two figures it can be seen that at *E*-region altitudes the current is mainly a Hall current. The shift in the ratio between the perpendicular conductivities is important in the study of ionospheric currents (*cf.* Chapter 5) and irregularities (Chapter 8).

Figure 4.3 also shows that except for the lowest part of the ionosphere the parallel conductivity, σ_{\parallel} , is much higher than the perpendicular components, and consequently $E_{\perp} \gg E_{\parallel}$. This means that at high latitudes, where the geomagnetic field lines are nearly vertical, the ionospheric layers are effectively coupled together, so that large-scale electric fields will spread in altitude without significant damping. Horizontal electric fields will thus become almost height independent.

Under such conditions it is meaningful to define a height-integrated linear current density as

$$\mathbf{J}_Z = \int \mathbf{j} dh. \quad (4.13)$$

Provided v_n is height independent or negligible, integration of Eq. (4.7) gives

$$\mathbf{J}_Z = \Sigma_P(\mathbf{E}_{\perp} + \mathbf{v}_n \times \mathbf{B}) + \Sigma_H \mathbf{B} \times (\mathbf{E}_{\perp} + \mathbf{v}_n \times \mathbf{B})/B, \quad (4.14)$$

where

$$\Sigma_P = \int \sigma_P dh, \text{ and } \Sigma_H = \int \sigma_H dh, \quad (4.15)$$

are known as the height-integrated conductivities.

It should be noted that the above only concerns DC conductivities. When it comes to small-scale electric fields and AC conductivities, the situation is quite different – drastic damping occurs for fields with horizontal scale of 1 km or less (Spreiter & Briggs 1961).

It is furthermore supposed that particle motions and currents are unlimited in all directions. In cases where currents are prevented from flowing in the directions associated with the primary electric field, secondary polarization fields can arise and cause an enhanced effective conductivity (*cf. e.g.* Boström 1972).

4.3. Drift motion of ionization irregularities

The drift velocities of electrons and ions, v_e and v_i , which are caused by the imposed electric field, will have a common component, v_d^0 , in the direction γ with the field (*cf.* Fig. 4.1). This velocity, which is called the neutral ionization drift, is perpendicular to both \mathbf{B} and \mathbf{j} . By use of Fig. 4.1 we find

$$v_d^0 = v_e \sin \alpha_i = \frac{E_{\perp}}{B} \sin \alpha_i \sin \alpha_e, \quad (4.16)$$

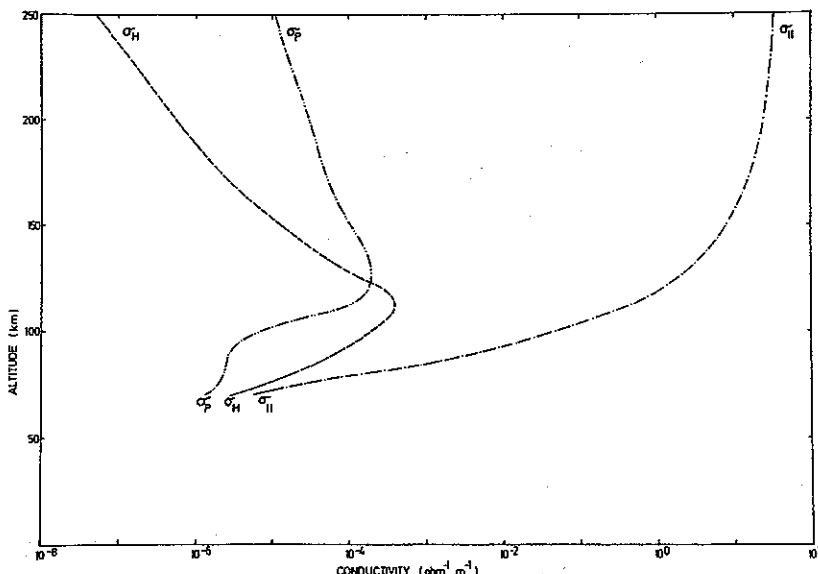


Fig. 4.3. Variation of the ionosphere conductivities with height for the mean ionospheric model used in Chapter 3.

and

$$\gamma = \frac{\pi}{2} - \alpha_e + \alpha_i. \quad (4.17)$$

Supposing that the driving Ampere force, $\mathbf{j} \times \mathbf{B}$, balances the collisional friction, one will obtain further

$$\mathbf{v}_d^0 = \frac{\mathbf{j} \times \mathbf{B}}{N_e(m_e \nu_e + m_i \nu_i)}. \quad (4.18)$$

Although this neutral ionization drift, which denotes the drift of an uniformly distributed ionization, has a well-defined physical meaning, it is not easily measured. Another quantity which is more readily traced, and therefore of more practical interest, is the motion of an ionization irregularity, where the boundaries are defined by gradients in N_e . This has been studied by, for example, Kato (1963, 1964), whose method will be used in the following.

The drift velocity of the irregularity can be composed of two parts, one perpendicular to \mathbf{j} corresponding to the neutral ionization drift, and one along \mathbf{j} . If displacement currents are neglected, $\nabla \times \mathbf{B} = \mu \mathbf{j}$ and consequently $\nabla \mathbf{j} = 0$. The variation of electron density due to the motion is then

$$\begin{aligned} \frac{\partial N_e}{\partial t} &= - \frac{\nabla \cdot (\mathbf{j} \times \mathbf{B})}{m_e \nu_e + m_i \nu_i} \\ &= - \frac{1}{m_e \nu_e + m_i \nu_i} [\mathbf{B} \cdot (\nabla \times \mathbf{j}) - \mu \mathbf{j}^2]. \end{aligned} \quad (4.19)$$

(Density variations due to production and loss processes, see Chapter 3, will in the following be neglected.)

If we only consider the motion in a plane perpendicular to \mathbf{B} the current is expressed as

$$\mathbf{j} = \sigma_P \mathbf{E} + \sigma_H \frac{\mathbf{B} \times \mathbf{E}}{B}. \quad (4.20)$$

From this one obtains

$$\begin{aligned} \mathbf{B} \cdot (\nabla \times \mathbf{j}) &= [\sigma_P (\mathbf{E} \times \mathbf{B}) + \sigma_H B \mathbf{E}] \cdot \frac{\nabla N_e}{N_e} \\ &\quad + \sigma_H B \nabla \cdot \mathbf{E} - \sigma_P B \cdot \frac{\partial \mathbf{B}}{\partial t}, \end{aligned} \quad (4.21)$$

using Maxwell's equation $\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t}$ together with vector identities (*cf. e.g. Rottman 1961*) and assuming σ_P and $\sigma_H \propto N_e$. Furthermore, from Eq. (4.20) one gets

$$\nabla \cdot \mathbf{E} = - \frac{\mathbf{j} \cdot \nabla N_e}{\sigma_P N_e} - \frac{\sigma_H}{\sigma_P} \frac{\mathbf{B}}{B} \cdot \frac{\partial \mathbf{B}}{\partial t}. \quad (4.22)$$

Using this in Eq. (4.21) we obtain

$$\mathbf{B} \cdot (\nabla \times \mathbf{j}) = \sigma_C \left[(\mathbf{E} \times \mathbf{B}) \cdot \frac{\nabla N_e}{N_e} - \mathbf{B} \cdot \frac{\partial \mathbf{B}}{\partial t} \right], \quad (4.23)$$

where $\sigma_C = \sigma_P + \frac{\sigma_H^2}{\sigma_P}$, also known as the Cowling conductivity.

Substitution of Eq. (4.23) into (4.19) then finally gives

$$\frac{\partial N_e}{\partial t} = K (\mathbf{E} \times \mathbf{B}) \cdot \nabla N_e + K N_e \left(\frac{1}{2} \frac{\partial B^2}{\partial t} + \frac{\mu \mathbf{j}^2}{\sigma_C} \right), \quad (4.24)$$

with

$$K = \frac{\sigma_C}{N_e(m_e \nu_e + m_i \nu_i)} = \frac{1}{B^2} \frac{\omega_{ce} \omega_{ci}}{\omega_{ce} \omega_{ci} + \nu_e \nu_i}. \quad (4.25)$$

In the derivation of the last form of K we have used

$$\sigma_C = \frac{e N_e}{B} \frac{\omega_{ce} \nu_i + \omega_{ci} \nu_e}{\omega_{ce} \omega_{ci} + \nu_e \nu_i},$$

which is obtained from algebraic manipulations after inserting the expressions for σ_P and σ_H (Eq. 4.9) in the equation defining σ_C .

In Eq. (4.24) the term containing \mathbf{j}^2 contributes to constant increase in N_e caused by the Ampere force due to the current and the induced magnetic field, *i.e.* a pinch effect (see *e.g. Singleton 1962*).

The $\frac{\partial B}{\partial t}$ - term is related to density variation by magnetic pressure. If we neglect these two terms, that is the same as assuming that the fields are electro- and magnetostatic, *i.e.* $\nabla \times \mathbf{E} = - \frac{\partial \mathbf{B}}{\partial t} = 0$ and $\mathbf{B} = \mathbf{B}_0$, which can be done at least under moderately disturbed conditions, Eq. (4.24) appears as

$$\frac{\partial N_e}{\partial t} = - \mathbf{v}_d \cdot \nabla N_e, \quad (4.26)$$

showing that the irregularity propagates with the velocity

$$\mathbf{v}_d = K (\mathbf{E} \times \mathbf{B}) = \frac{\omega_{ce} \omega_{ci}}{\omega_{ce} \omega_{ci} + \nu_e \nu_i} \frac{\mathbf{E} \times \mathbf{B}}{B^2}. \quad (4.27)$$

It should be noted that \mathbf{E} here is the total electric field, and that it thus also contains terms due to polarization fields in the irregularity. This component depends on $\frac{\nabla N_e}{N_e}$ and \mathbf{j} (cf. first term in Eq. (4.22)) and will highly modify the motion of the irregularities and can also cause deformations of the shapes (cf. e.g. Kato 1963, 1964).

Using the first version of K (Eq. (4.25)), v_d appears as

$$\begin{aligned} v_d &= \frac{\mathbf{j} \times \mathbf{B} - \frac{\sigma_H}{\sigma_P} B \left(\sigma_P \mathbf{E} + \sigma_H \frac{\mathbf{B} \times \mathbf{E}}{B} \right)}{N_e(m_e v_e + m_i v_i)} \\ &= v_d^0 + \left(\frac{\sigma_H}{\sigma_P} B \right) \frac{\mathbf{j}}{N_e(m_e v_e + m_i v_i)}. \quad (4.28) \end{aligned}$$

The drift velocity of an ionospheric irregularity in a plane perpendicular to the magnetic field will thus consist of two components. The first corresponds to the motion of the ambient ionization, v_d^0 . The second is in direction antiparallel to the current and in magnitude $\frac{\sigma_H}{\sigma_P}$ times v_d^0 , so that for the regions where $\frac{\sigma_H}{\sigma_P} \gg 1$ the last term will dominate the drift. While the v_d^0 -component corresponds to a real body motion, the other velocity, $\frac{\sigma_H}{\sigma_P} v_d^0$, is associated with a wave motion. This will be further discussed in Chapter 8.

5. CURRENT SYSTEMS AND LARGE-SCALE ELECTRIC FIELDS IN THE HIGH LATITUDE IONOSPHERE

5.1. Introduction

The presence of ionospheric currents was postulated long before any measurements of electric fields were made. We will therefore first briefly review the present picture of the auroral current systems. In the last, and main part of the chapter, large-scale electric fields will be considered. We will summarize the present knowledge of high latitude E -fields, and present the technique for direct probe measurements of electric fields. Finally, the DC fields observed during two specific events, and the relation between electric

fields and optical aurora, will be examined in more detail.

5.2. Ionospheric current systems

Historically, interest in this topic was raised from the observations of geomagnetic phenomena. Short-time variations in the earth's magnetic field were discovered in 1722 by George Graham (Graham 1724) and later verified in 1740 by Andreas Celcius (Celcius 1741). In 1741 Celcius also noted that there was a close connection between aurora and magnetic disturbance. The first to suggest that atmospheric currents might produce these variations was Carl F. Gauss, in his *Allgemeine Theori des Erdmagnetismus* in 1839, even though he could not understand how such currents could possibly exist. A conducting layer in the upper atmosphere was predicted by Balfour Stewart in 1882 (20 years before the experimental radio 'discovery' of the ionosphere), and he ascribed the daily magnetic variations to electric currents flowing in this layer.

Studying high latitude magnetic perturbations, Kristian Birkeland in 1908 proposed a *system* of currents in the upper atmosphere to explain the observed variations in the magnetic field during 'polar elementary storms' (now known as polar magnetic substorms). Combining magnetic data from several stations he inferred a horizontal distribution of currents causing the perturbations. The complete system consisted of horizontal currents in the upper atmosphere supplied by vertical (geomagnetically field aligned) currents from the magnetosphere. He was, however, fully aware that this system was only one of many possible models, and that it did not necessarily represent the actual current flow (Birkeland 1908, pp. 43-44).

In 1935, Chapman launched a competitive current system to explain the magnetic disturbances. This system, which was taken up by Vestine (Vestine & Chapman 1938) and others, grew to dominate the picture of ionospheric currents in the coming three decades. (Schematic illustrations of the Birkeland and Chapman currents systems are shown in Fig. 5.1.) Chapman's current pattern is entirely two-dimensional and is composed of a strong bimodal current, electrojet, along the

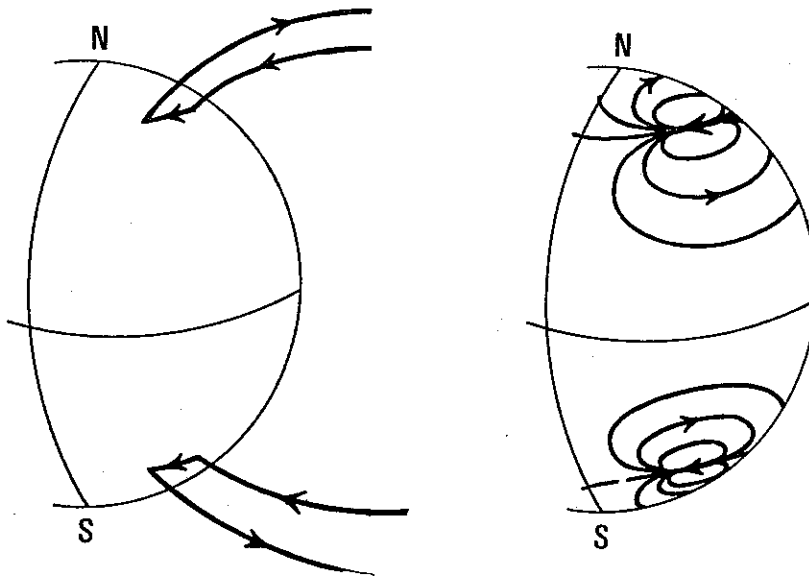


Fig. 5.1. Schematic illustration of the Birkeland (left) and Chapman (right) current systems for a polar magnetic substorm (after Fukushima 1969).

auroral oval, tied together with return currents at higher and lower latitudes. The magnetic effects as observed on the ground at auroral and middle latitudes, are, however, the same for both systems (Fukushima 1969), and for several reasons the Chapman-Vestine view was preferred by most workers in the field rather than the Birkeland system, which was mainly advocated by Alfvén (e.g. Alfvén 1939, 1940, 1950).

In recent years, with observations from rockets and satellites, several independent investigations have, however, provided experimental evidence for field-aligned currents (Zmuda *et al.* 1966; Hoffman & Evans 1968; Armstrong & Zmuda 1970; Cloutier *et al.* 1970; Park & Cloutier 1971), and three-dimensional current-systems based on Birkeland's ideas were again brought forward (Boström 1964, 1968; Cunnings & Dessler 1968). (In 1972 it was suggested by the International Geophysical Union that field-aligned currents in the ionosphere and magnetosphere should be called Birkeland currents.) Figure 5.2 shows a current configuration deduced by Park & Cloutier (*loc. cit.*) from rocket observations.

Besides the change from a two-dimensional to a three-dimensional model, the picture has also changed from a completely static pattern of the kind shown in Fig. 5.3, where the earth is rotating under a fixed current system, to a more dynamic and variable pattern. We will in the following tie together in a descriptive form some of the latest

ideas, in an attempt to give an up-to-date picture of the auroral electrojet. (For a definition of the various terms used in connection with the substorm concept, see e.g. Davis 1972.)

The direction of the electrojet changes, not only

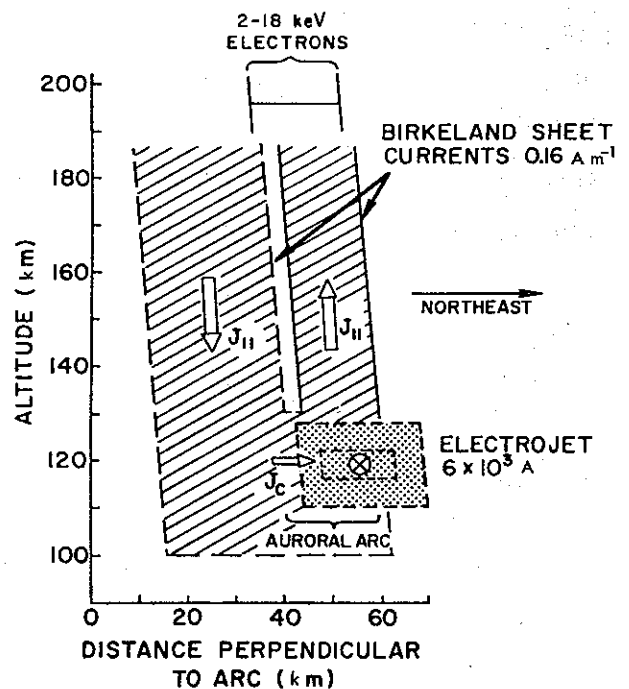


Fig. 5.2. Current system deduced by Park & Cloutier (1971) from rocket measurements of particles and magnetic fields. The complete current configuration consists of an east-west electrojet, two Birkeland sheet currents, $J_{||}$, at the edges of an auroral arc, and a closing current, J_c , between the Birkeland currents.

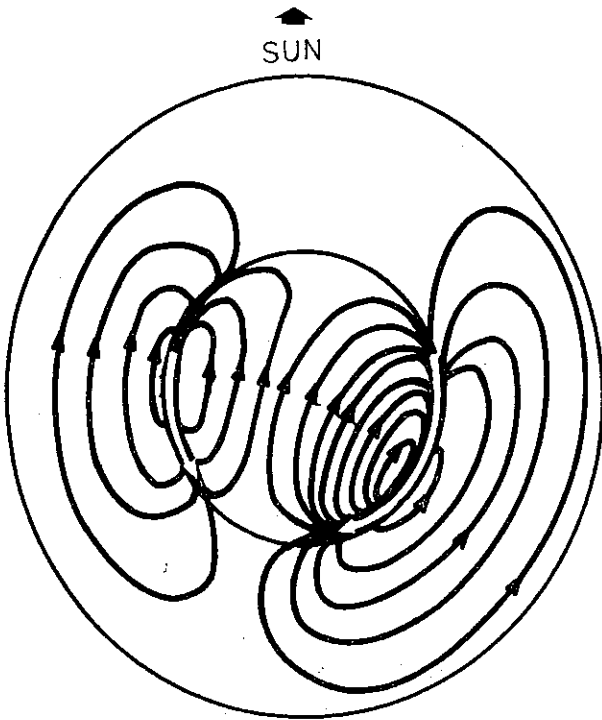


Fig. 5.3. Old model of current system for polar magnetic substorm, view from above the magnetic north pole.

with local time, as in the old picture, but also with the development of the substorm. During the quiet phase there is strong evidence of a bimodal electrojet pattern, with eastward currents in the afternoon and evening sectors, and westward currents in the morning sector, as in the old picture. This configuration will also exist during the growth phase of the storm. The picture in the midnight sector is, however, particularly uncertain, and the mapping of the flow in this discontinuity region, where a reversal of the electrojet takes place, is quite ambiguous and variable. Since Harang (1946) was the first to draw attention to the structure of the currents in this sector, it has been suggested (Heppner 1972a) that the region should be called the 'Harang discontinuity'.

At the start of the expansive phase (breakup, $t = 0$ in Akasofu's substorm picture (Akasofu 1968)), the picture apparently changes to a unimodal westward current. Figure 5.4 gives a picture of the configuration during this phase.

The recovery phase will also bring about a recovery of the auroral electrojet. There is a

widespread current in the morning sector, and a growth of the eastward current along the auroral zone in the afternoon and evening sectors, with a gradual transition to a quiet-phase picture.

Measurements of drift motions of artificial ion clouds (see Sect. 5.5) in the magnetosphere (Haerendel & Lüst 1970) and polar-cap ionosphere (Westcott *et al.* 1970) have produced strong evidence against the existence of return currents over the polar cap. Further analysis of measurements of electric fields and magnetic perturbations led Heppner *et al.* (1971a,b) to the conclusion that practically all polar-cap horizontal disturbances in the magnetic field had to be explained in terms of non-ionospheric currents. A mid-latitude ionospheric current to close the auroral electrojet was ruled out for the reasons that both the night-time electron density and electric field were too low to support a current of sufficient intensity. The model proposed for a continuation of the auroral electrojet is shown in Fig. 5.5. Field-aligned currents complete the electrojet circuit, and their general ionospheric pattern of current flow gives an inward current in the morning sector and an outward current in the Harang discontinuity region.

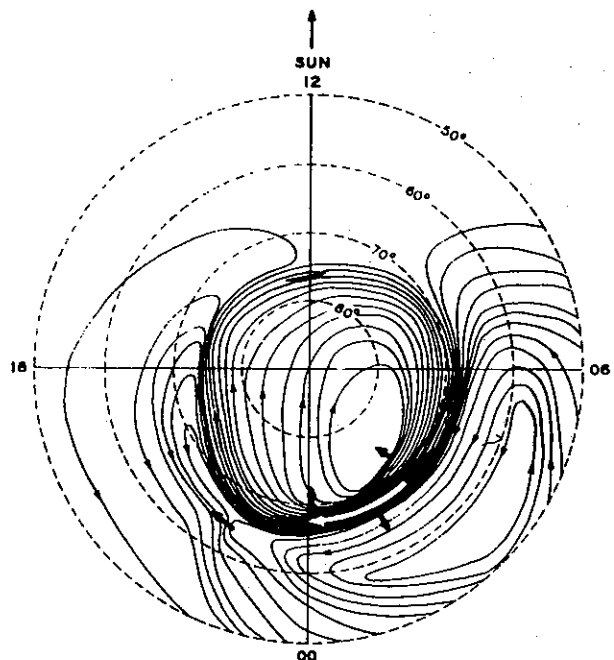


Fig. 5.4. New model of current system during the expansive phase of a substorm, view from above magnetic north pole (after Akasofu 1966).

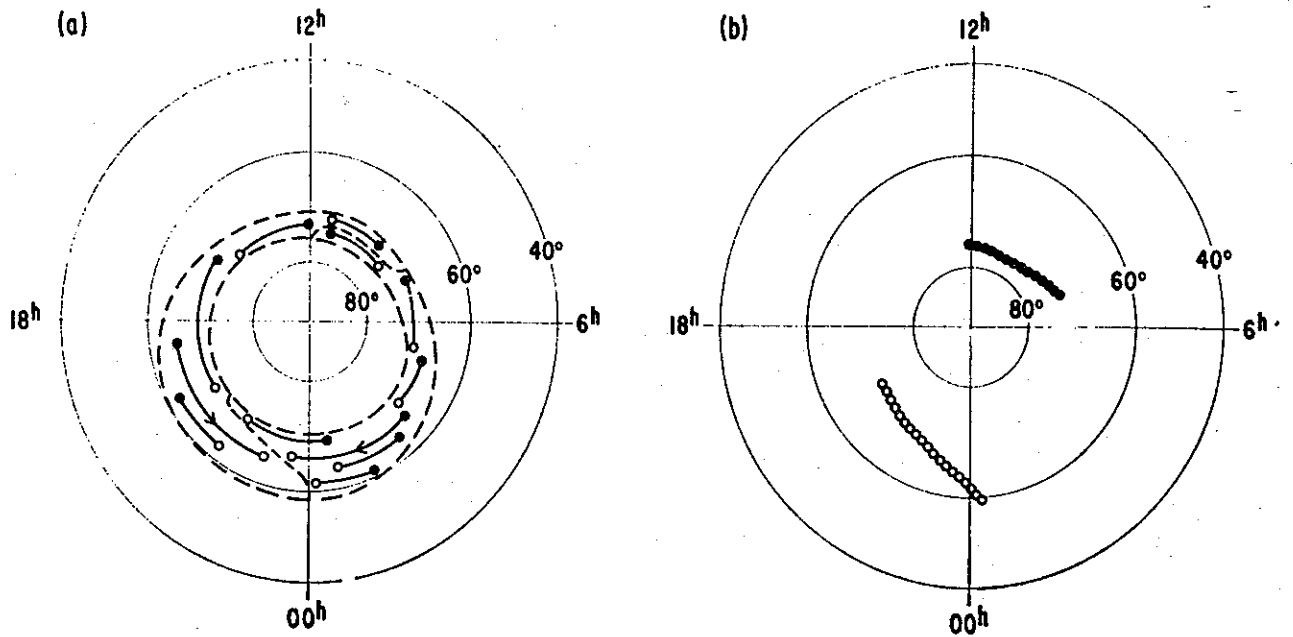
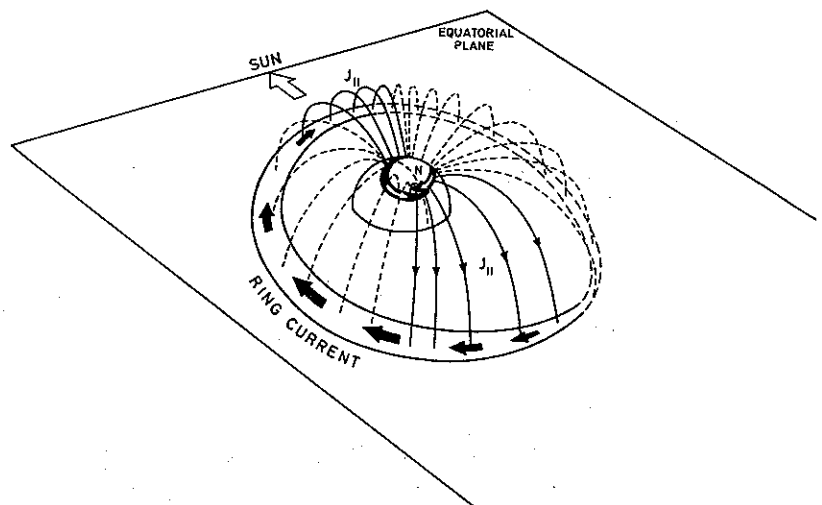


Fig. 5.5. Schematic illustration of electrojet Hall currents terminated in Birkeland currents (part (a)). Part (b) shows the net distribution of Birkeland currents, resulting in sheet configurations of net field-aligned current. Open and filled circles indicate currents 'out of' and 'into' the ionosphere, respectively (after Heppner *et al.* 1971a).

Figures 5.6 and 5.7 will then give a representative picture of the present model of the auroral current system. During the quiet, growth, and the last part of the recovery phase of the substorm, the electrojet is bimodal, and is continued in field-aligned currents, which give an inward net current flow in the morning sector and an outward flow in the Harang discontinuity (Fig. 5.6). The closure of the system in the magnetospheric equatorial plane is consistent with an asymmetric

ring current. The ionospheric polar cap and mid-latitude return circuits are weak. In the expansive phase and the first part of the recovery phase, the current system consists of an unimodal westward electrojet (Fig. 5.7). Field-aligned currents, flowing into the ionosphere from the morning sector and out to the magnetosphere in the evening sector, establish a connection between the magnetospheric ring current and the ionospheric system.

Fig. 5.6. Schematic three dimensional model of a current system for the quiet, growth, and recovery phase of a substorm. The system consists of a bimodal electrojet, linked to Birkeland currents giving a net inward current flow in the morning sector and an outward flow in the midnight sector. The closure of the Birkeland currents is in an asymmetric ring current in the magnetosphere. Fully drawn lines represent regions with major net currents.



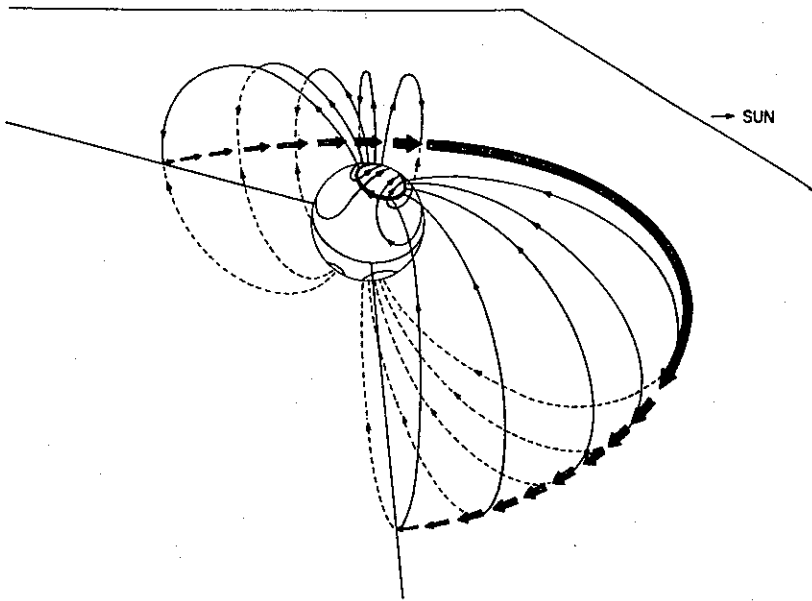


Fig. 5.7. Three dimensional current system for the expansive phase of the substorm, consisting of an unimodal electrojet, Birkeland sheet currents and an asymmetric ring current (after Akasofu & Meng 1969).

5.3. Sources of DC electric fields

The driving forces of the auroral electrojet will be electric fields of global structure. The sources of the fields will primarily be found in the magnetosphere, but one must also consider fields of more local character.

Electric fields in the magnetosphere can be produced by several mechanisms. Magnetic field curvature and gradients cause the electrons and ions to move with slightly different velocities, which in turn will bring them into different regions. The charge separation will produce an electric field, and it has been suggested that this should be the primary mechanism for magnetospheric fields (Fejer 1961; Kern 1962). In another theory, introduced by Alfvén (1960) and later taken up and modified by several others (e.g. Karlson 1963; Block 1966), the magnetospheric electric field was supposed to be the field induced within a magnetized solar plasma beam.

These mechanisms will certainly have influence on the electric fields in the magnetosphere and ionosphere. However, recent investigations of e.g. auroral motions (e.g. Davis 1971), dynamics of the plasmasphere (e.g. Carpenter 1970), and electric fields (e.g. Kaufman & Gurnett 1971) have called new attention to convection processes proposed by Axford & Hines (1961) and Dungey (1961) as the principal source of magnetospheric electric fields. Energy and momentum from the solar wind may be transferred to the magneto-

sphere by a viscous-like interaction at the boundary of the magnetosphere. This energy induction will cause a convection of the magnetospheric plasma, i.e. the low-energy plasma experiences a large-scale circulation in the geomagnetic field, as shown in Fig. 5.8a. Another possible

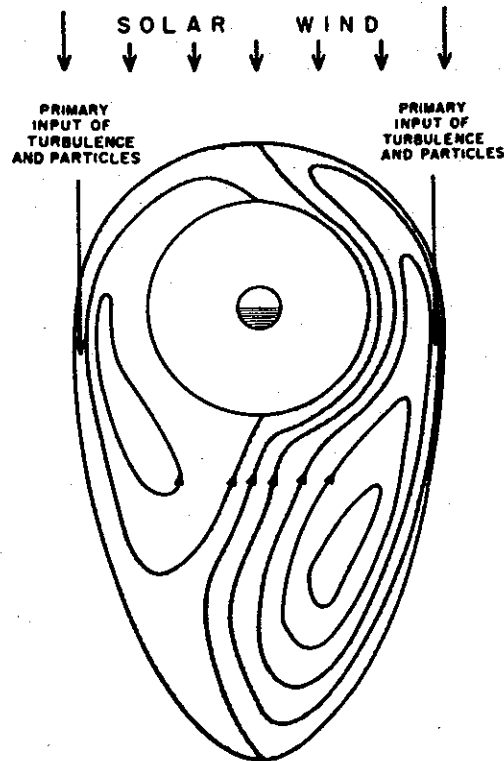


Fig. 5.8.a. Model of magnetospheric convection patterns as originally suggested by Axford & Hines (1961).

energy input is through reconnection of magnetic field lines in the tail section as illustrated in Fig. 5.8b (cf. Dungey 1961). However, regardless of the type of interaction which takes place, the result will be a convection of the plasma inside the magnetopause. This motion of the charged particles will induce an electric field (cf. Eq. (4.1)) $\mathbf{E} = -\mathbf{v} \times \mathbf{B}$, and a connected electrical potential, ϕ . If $\partial/\partial t = 0$, i.e. steady motion, we will have $\mathbf{E} = -\nabla\phi$ and hence the streamlines of the convective motion and the magnetic field lines will be equipotential. However, the behaviour of the magnetosphere during a substorm has clearly demonstrated that the convection is not steady and that $\partial\mathbf{B}/\partial t \neq 0$, and instead of the simple relationship between field and potential we will have $\mathbf{E} = -\nabla\phi - \frac{\partial\mathbf{A}}{\partial t}$, where \mathbf{A} is the magnetic vector potential defined by $\mathbf{B} = \nabla \times \mathbf{A}$. However, Eq. (4.1) is still valid, so that \mathbf{E} will remain perpendicular to \mathbf{v} and \mathbf{B} .

Without going any further into the processes of generation, in the following we will pay attention to the behaviour of the electric field.

Due to the high parallel conductivity (cf. Sect. 4.2) the electric field will propagate along the magnetic lines of force. Although the large scale fields are little attenuated in this propagation (cf. e.g. Farley 1959), the mapping onto the ionosphere cannot, as is frequently done, be assumed as perfect. It has for instance been shown (Mozier 1970) that the direction of the field will not be exactly preserved, due to different attenuation of the different components. Furthermore, finite potential drops will occur along the field lines. If the voltage drop is different along different field lines, this will give horizontal potential differences in the ionosphere not due to magnetospheric fields. Any convection pattern generated outside the ionosphere can also be strongly affected by regions of high conductivity (e.g. an auroral arc). The magnetospheric plasma may tend to flow around and thus avoid such regions (Baker & Hammel 1962). An alternative way to regard this is to consider the E -field source as a constant current generator (Aggson 1969). A high conductivity region in the ionosphere will load the generator, thus causing a voltage drop.

Differences in propagation conditions and con-

ductivities will thus disturb the image of magnetospheric fields exposed in the ionosphere. On a large scale, however, the distribution of electric fields in the ionosphere and in the equatorial plane of the magnetosphere must be related.

5.4. Characteristics of large-scale electric fields

Up to about 1960 the interest in ionospheric electric fields was very low, and no serious attempt was made to accomplish direct field measurements. The main reason for this was that infinite conductivity was generally assumed along the magnetic field lines. The fields could therefore be obtained straightforward from measurements of \mathbf{B} . This attitude has completely changed, and the

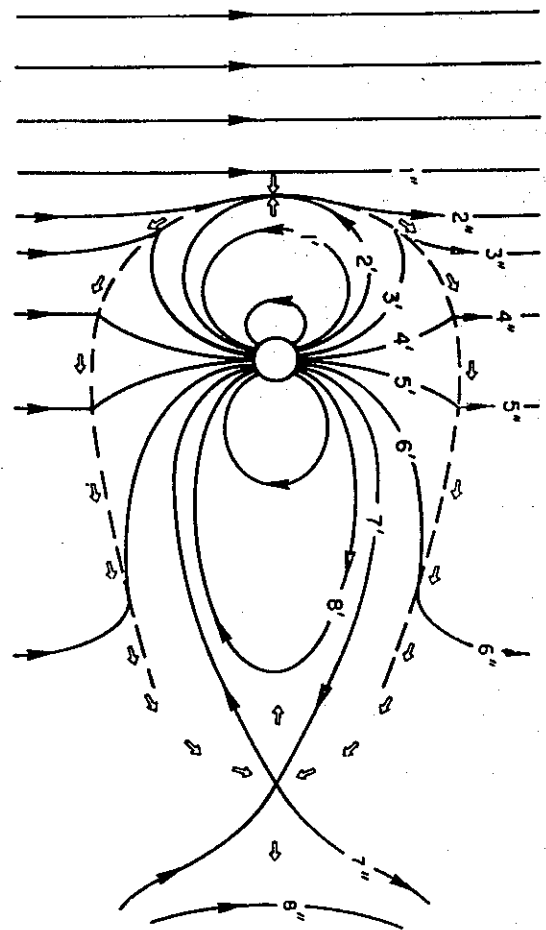


Fig. 5.8.b. Interaction between interplanetary and geomagnetic field and the resulting plasma flow (open arrows) as proposed by Dungey (1961). Numbers indicate motion of individual field lines with the motion progressing toward higher numbers. Reconnection occurs at the contacts 2'-2'' and 7'-7'' (Levy *et al.* 1964).

electric field is now regarded as an important and necessary parameter in the understanding of the physics of the ionosphere. The last decade's direct and indirect determinations of ionospheric *E*-fields have increased our knowledge of the nature of fields considerably. However, when combining the various groups of information, the resulting picture is far from unambiguous. In this section we will review its main aspects.

There is general agreement concerning the magnitude of the fields. Typical values in the auroral oval are 30–50 mV m⁻¹, but variations from less than 10 to 150 mV m⁻¹ are not unusual. The field can, however, undergo large variations in magnitude and direction over short distances and times (*cf.* Fig. 5.17). In most cases there seems to be no systematic variation with altitude of fields perpendicular to **B** (Potter 1970; Maynard 1972), but *e.g.* Mozer & Fahlson (1970) have published results contradicting this.

The field has been shown (*e.g.* Wescott *et al.* 1969; Fahlson *et al.* 1971) to be predominantly poleward during positive bays and equatorward during negative bays (Figs. 5.9 and 5.13). This confirms that the currents causing the magnetic perturbations at auroral latitudes are mainly Hall currents. It has, however, been clearly demonstrated (*e.g.* Haerendel *et al.* 1969; Wescott *et al.* 1969) that there is no simple relation between the magnitudes of the electric field and the magnetic perturbation. This means that the conductivities, which are as important for the determination of the current as the field itself, are subjected to great variations. As the surface magnetic observatory will only see integrated effects of the regional ionospheric current, it is clear that the possibility

of deducing information about the fields from observed magnetic disturbances is very limited.

The electric field in the midnight sector will, in the same way as the current configuration, undergo large variations. The instantaneous form and location of this discontinuity region, where convection patterns, motion of auroral forms, and the direction and magnitude of the fields are rather confused, are frequently changing (Heppner 1972*a*).

The question of what happens to the electric field near and inside an auroral form has also been subject to discrepancies. Aggson (1969) deduced from probe experiments that the field in an arc was greatly reduced, and this seemed to be confirmed by other experiments (Wescott *et al.* 1969; Potter & Cahill 1969). However, Mozer & Fahlson (1970) found cases where the field on average remains constant when crossing the boundary of an arc, but large amplitude fluctuations can be present, both at the boundary and inside the arc. Other measurements have given examples where the field stays rather constant in the transition of the aurora (Ungstrup, *priv. com.*). Thus the results are controversial, and the situation seems to change from one case to the other. It should, however, be noted that none of these authors made *in situ* measurements of the optical emissions, but used particle detectors and television cameras (Mozer & Fahlson), triangulation from several ground camera sites (Wescott *et al.*), particle detectors (Potter & Cahill), or particle detectors, ground photometers and cameras (Aggson, Ungstrup) to identify the location of the optical forms. (We will return to this subject later in this chapter.)

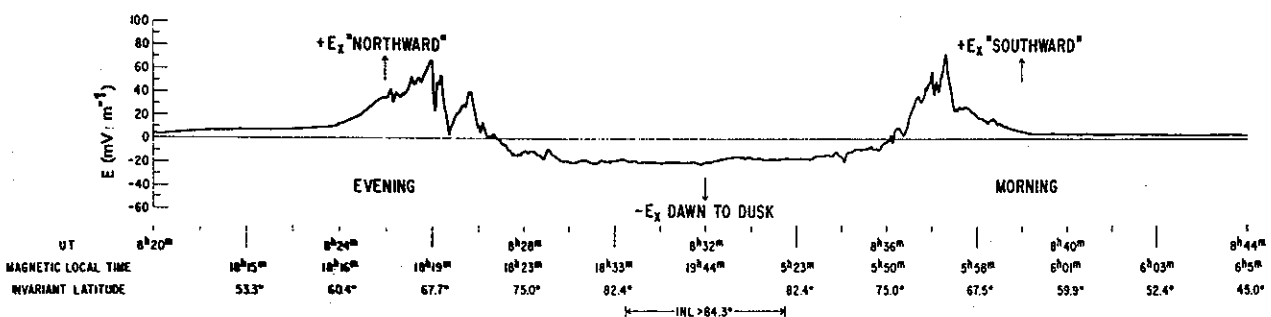


Fig. 5.9. A northern polar pass of OGO-6, showing the electric field at auroral latitudes and over the polar cap (after Heppner 1972*b*).

The existence of electric fields parallel to \mathbf{B} in the auroral zone is another controversial question. Important as such fields would be for particle acceleration and precipitation, they have been suggested by many theorists (*cf. e.g.* Alfvén & Fälthammar 1963; Persson 1967). Probe measurements by Mozer & Bruston (1967, later re-examined by Mozer & Fahleson 1970) seemed to verify the existence of a parallel field. However, their value of $\approx 20 \text{ mV m}^{-1}$ seems to be too high, compared to theoretical values. Mozer and Fahleson also found an increase with altitude of the perpendicular field, which maintains $\nabla \times \mathbf{E} = 0$ required in the steady-state condition assumed in their derivation.

On the other hand, several other workers (*e.g.* Mende 1968; Wescott *et al.* 1969; Potter 1970) using different observation techniques, deduced that no parallel fields of any importance ($< 60 \mu\text{V m}^{-1}$) existed. Some of the assumptions made in the pro-parallel-field papers were also thrown into doubt.

Recent probe measurements from the satellite Injun-5 have, however, shown that weak parallel fields may be present (Cauffman & Gurnett 1971). Hultqvist (1971) has also obtained new experimental evidence for the existence of parallel fields in observations of field-aligned anisotropy in ion pitch-angle distribution in the auroral zone and over the polar cap.

Thus the dispute on parallel fields continues. However, the bulk of the evidence seems to support the statement that parallel fields do exist, but that they are generally markedly less (order of magnitude) than the perpendicular component. In addition, electrostatic waves and short duration electric fields of much greater amplitude can be found.

Measurements of electric fields in the polar cap have led to a revised conception of the current configuration in these regions (*cf.* Sect. 5.2). From release experiments (Wescott *et al.* 1970) and satellite data from OGO-6 (Heppner 1972*b*) it was found that in the polar cap the convection velocity was away from the sun and the resulting

Table 5.1. Characteristics of high latitude electric fields

PROPERTY	CONCLUSION	KEY REFERENCE
Magnitude	30–50 mV m^{-1} (typical) 10–150 mV m^{-1} (freq. obs.)	Wescott <i>et al.</i> 1970
Amplitude variation	Variable over short dist. and time No systematic variation with altitude	Haerendel <i>et al.</i> 1969 Potter 1970, pro; Mozer & Fahleson 1970, contra.
Auroral electrojets	Hall currents	Wescott <i>et al.</i> 1969.
Relation \mathbf{E} , $\Delta\mathbf{B}$	Complex. Conductivities important	Haerendel <i>et al.</i> 1969.
E in aurora	Reduced in magnitude inside forms	Aggson 1969, pro; Mozer & Fahleson 1970, contra.
Parallel fields	Exist, but less than perp. fields	Mozer & Fahleson 1970, pro; Potter 1970, contra.
Polar cap fields	Directed dawn-dusk Polar cap currents: field aligned	Heppner 1972. Heppner <i>et al.</i> 1971.

electric field pointing toward the evening sector, thus giving rise to a field reversal at the poleward boundary of the auroral zone (Fig. 5.9). The observed ground magnetic variations excluded Hall or Pedersen currents as a possible explanation, and a field-aligned current system completing the electrojet circuit is proposed (Sect. 5.2).

The polar-cap field was found to be quite uniform in space and time (Heppner *loc. cit.*), but small scale changes are very evident (Maynard & Heppner 1970). The magnitude of the polar-cap field is typically $20\text{--}40\text{ mV m}^{-1}$.

Conclusions from the measurements of high-latitude electric fields in the ionosphere are summarized in Table 5.1. Improved experimental techniques have provided new information about the magnitude and structure of the fields and their associated currents. This in turn has given a better understanding of the origin of the electric fields. However, important questions, such as the existence of fields parallel to \mathbf{B} and the configuration in the noon and midnight sectors, are still unanswered.

5.5. Experimental techniques for *in situ* measurements of electric fields

In situ measurements of electric fields are cluttered by several difficulties and have historically lagged behind the theories for fields and current systems in the ionosphere. Indirect determinations, mainly based on measurements of drift velocities, magnetic perturbations, and particle parameters, suffer from a lack of exact information about other properties involved in the field calculations.

During the last few years two different techniques, the release technique and the double floating probe technique, have, however, proved capable of giving dependable results.

The principle of the *release technique* is that an artificial cloud, usually barium, is injected into the ionosphere. This cloud will be partly ionized by photon-ionization. While the neutral cloud drifts along with the wind system, the ionized cloud will be under the additional influence of electric and magnetic fields (*cf.* Sect. 4.3). By determining the velocity of the Ba^+ -cloud relative to the neutral constituents it is possible to deduce

information about the direction and magnitude of the electric field.

The technique has proved to be very useful for studying long-time variations of DC fields in the ionosphere and magnetosphere. However, the requirement of twilight conditions (sunlight for photon-ionization, but dark enough for optical tracking), severely restricts local time coverage. The presence of the ion cloud may also disturb the medium and give rise to plasma instabilities. (For details and a theoretical background for such measurements, see, for example, Haerendel *et al.* 1967.)

The electric field measurements performed in this investigation were based on the *double floating probe technique*, and in the following we will concentrate on the principles behind this kind of observation.

5.5.1. Probe design. Intuitively one has the feeling that the most straightforward way to accomplish a determination of electric field strengths is to establish two reference points in the plasma and measure the voltage between these. As the potential of the spacecraft is poorly defined, a monopole using the vehicle itself as reference point is generally unsuitable. A monopole will also be unduly sensitive to technological interference of various kinds.

The electric field can in principle be calculated from the voltage measured between two (or more) electrodes in contact with the plasma, if the geometry of the electrodes is known. There are, however, several practical difficulties involved in these measurements. A conducting body inserted into a plasma will in most cases attain a potential different from that of the surrounding plasma. This causes an attraction of charged particles with one sign, while particles with opposite charge will be repelled. Thus, a layer will be established near the probe where charge neutrality does not exist. The thickness of this plasma sheath will depend on the shape, surface, and dimension of the probe, the Debye length, and a function monotonically increasing with the ratio probe potential to plasma thermal energy (Storey 1963). The voltage drop in the sheath is very difficult to determine both theoretically and in plasma laboratories. It will

also most probably be different for differently shaped and oriented probes. A fundamental rule for probe measurements is therefore that the sensors should be equally shaped, that they should be made from the same material and that their electrical connections should be identical. Another requirement is that the electrical mid-points of the probes should be far away from each other (several metres). This provides a long baseline for the measurements, giving a potential difference which is large compared to voltages from disturbing effects, and removes the active parts of the antennas away from interference sources and regions screened off by the spacecraft.

The types of probes used for electric field measurements can be designated either as 'long cylindrical booms' or 'spherical antennas'. The first group can be long metal rods, either plain metal, insulated, or partly insulated, while the second type of probes consists of small metal spheres supported by insulating rods. The probes used in our measurements were of the first kind, and the following considerations will therefore be limited to the long cylindrical boom technique.

With the *STEM* technique (*STEM*: Storable Tubular Extendible Member) originally developed for the Alouette satellites (*cf. e.g.* Mar & Garrett 1969), it is possible to attain a large separation between the electrical centres of an antenna system consisting of two cylindrical booms. As the antenna elements in most cases are erected from the vehicle itself, the difficulty with a long connection line between the electronics and antenna is avoided. The disadvantages with this probe system arise mainly from the extension of the contact area between the electrode and the plasma. One may thus have variations over the dimension of the probes which make the geometry of the system somewhat uncertain. The steady electric field will also have influence on the AC properties of the antenna, creating a displacement between the electrical and geometrical centres of the antenna (Storey 1963). For a symmetrical dipole system of two similar elements extended in opposite directions, the displacement of the electrical centres should in theory be the same so that the total effective length will remain constant. But, even so, the probe system may be unbalanced

with respect to the spacecraft body, making the measurements more exposed to technological interference. Another circumstance which can affect the signal from the booms is that a part of the antenna will be in a region perturbed by the spacecraft body (*cf.* Appendix I).

Both these effects can be eliminated by coating the antenna elements with insulating material. This will, however, introduce the defect that the antenna impedance will be high at extremely low frequencies and also make the probes inapplicable for DC measurements. Where DC and low frequency response is wanted, one has therefore to adopt a compromise solution where part of the antenna element is left uninsulated.

5.5.2. Probe potential and induced potential difference. The potential of the probe with respect to the surrounding plasma will govern the currents in the plasmasheath. These currents, J_e : electron current, J_i : ion current and J_{ph} : current caused by photon-emission of electrons from the electrode, must under stationary conditions satisfy the current balance equation

$$J_e + J_i + J_{ph} = J. \quad (5.1.a)$$

J will then be the current flowing away from the probe (to the measuring apparatus). Other charging currents such as secondary emissions have been omitted.

When no current is drawn from the probe it will attain its 'floating potential', φ_o , and the current balance will be

$$J_e(\varphi_o) = J_i(\varphi_o) + J_{ph}(\varphi_o). \quad (5.1.b)$$

If a Maxwellian velocity distribution is assumed, the electron and ion currents to a surface element dS of a cylindrical probe will be for retarding potentials (*cf.* Mott-Smith & Langmuir 1926)

$$dJ_k = j_{ko} \exp \left[\frac{-e\varphi}{\kappa T_k} \right] dS, \quad (5.2.a)$$

while the expression for accelerating voltages appears as

$$dJ_k = j_{ko} \left\{ \frac{b}{a} \left[1 - \operatorname{erf} \left(\frac{a^2 e \varphi}{(b^2 - a^2) \kappa T_k} \right)^{1/2} \right] + \operatorname{erf} \left(\frac{b^2 e \varphi}{(b^2 - a^2) \kappa T_k} \right) \right\} dS. \quad (5.2.b)$$

The function $\text{erf}(x)$, the error function, is defined by

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty \exp(-y^2) dy.$$

b and a are the sheath and probe radii, respectively, and j_{ko} denotes the surface current density at $\varphi = 0$. This is the same as the random current density in the gas, and can thus be expressed as a function of particle density and mean thermal velocity (*cf.* Sect. 3.5.3),

$$j_{ko} = \varepsilon_k \frac{1}{4} e N_k \langle v_k \rangle = \varepsilon_k \frac{1}{4} e N_k \left(\frac{8\kappa T_k}{\pi m_k} \right)^{1/2}. \quad (5.3)$$

$\varepsilon_k = +1$ for ions (regarded positively charged) and -1 for electrons.

As $j_{eo} \gg j_{io}$ it is necessary to have $\varphi < 0$ to establish the current balance, except for very high and very low altitude where j_{ph} may dominate over j_e (Fahleson 1967). In the following only negative probe potentials will therefore be considered. Under the conditions existing in the ionosphere the ion current density will not be much changed from j_{io} (Storey 1963) so that the currents can be expressed as

$$\begin{aligned} dJ_e &= j_{eo} \exp\left[\frac{-e\varphi}{\kappa T_e}\right] dS, \\ dJ_i &= j_{io} dS, \\ dJ_{ph} &= \gamma j_{pho} dS. \end{aligned} \quad (5.4)$$

The photon current density will be given by an expression similar to Eq. (5.3), and γ is a geometrical factor (< 1) which is mainly determined by the position of the sun with respect to the probe. The integration for J_{ph} goes over the illuminated probe surface.

The probe potential can now be found by combining Eqs. (5.1.b) and (5.4) which gives

$$\varphi_o \approx -\frac{\kappa T_e}{e} \ln \frac{j_{eo}}{j_{io} + j_{pho}}. \quad (5.5)$$

Estimating the value of φ_o in a night-time rocket where $j_{pho} = 0$ for the simple case $T_e = T_i$, $N_i = N_e$ and $m_i = 30 u$ one gets the rough value $\varphi_o \approx -5.4 \cdot \frac{\kappa T_e}{e}$. In general one can expect a floating

potential of the order $\frac{\kappa T}{e}$. However, φ_o will vary with the plasma parameters, a fact which must be taken into account in connection with measurements from rockets and satellites traversing regions with different properties.

If a potential gradient over the probe dimension exists in the plasma, or if the probe is moving in a magnetic field, the current balance will not necessarily be fulfilled for a small surface element. However, in equilibrium the total current to the probe must also in this case be zero, and

$$\begin{aligned} \int j_e(r, \varphi_o) dS \\ = \int j_i(r, \varphi_o) dS + \gamma \int j_{ph}(r, \varphi_o) dS. \end{aligned} \quad (5.6)$$

The probe system considered in the following will consist of two cylindrical booms, A and B , symmetrically orientated with respect to an origin. The boom radius is a , and the active surface extends from d_1 to d_2 . The electric field in the plasma is \mathbf{E} and the probes are moving with a velocity \mathbf{v} . If the potential of the booms with respect to the plasma potential at the origin is $\varphi_{A, B}$, the expression for the electron current through a surface element at distance r will for probe A be

$$\begin{aligned} dJ_e(r)_A &= j_{eoA} \\ &\times \exp\left[-\frac{e}{\kappa T_{eA}} (\varphi_A + (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{r})\right] dS, \end{aligned} \quad (5.7)$$

and similar for probe B (*cf. e.g.* Aggson *et al.* 1965).

Substituting this on the left side of Eq. (5.6) gives

$$\begin{aligned} 2\pi a j_{eoA} \int_{d_1}^{d_2} \exp\left[-\frac{e}{\kappa T_{eA}} (\varphi_A + (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{r})\right] dr \\ = \int_{d_1}^{d_2} j_{iA} dS + \gamma \int_{d_1}^{d_2} j_{phA} dS, \end{aligned} \quad (5.8)$$

and by integration

$$2\pi a j_{eoA} \exp\left[-\frac{e\varphi_A}{\kappa T_{eA}}\right] \left(-\frac{\kappa T_{eA}}{e(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \hat{\mathbf{r}}}\right)$$

$$\times \left\{ \exp \left[\frac{-e}{\kappa T_{eA}} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_2 \right] - \exp \left[\frac{-e}{\kappa T_{eA}} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_1 \right] \right\} = J_{IA} + J_{pHA}. \quad (5.9)$$

Rearranging this gives the following expression for the probe potential

$$\varphi_A = -\frac{\kappa T_{eA}}{e} \ln \left\{ \frac{\kappa T_{eA}}{e} \frac{2\pi a j_{e0A} \left\{ \exp \left[\frac{-e}{\kappa T_{eA}} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_1 \right] - \exp \left[\frac{-e}{\kappa T_{eA}} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_2 \right] \right\}}{(\mathbf{E} + \mathbf{v} \times \mathbf{B}) (J_{IA} + J_{pHA})} \right\}. \quad (5.10)$$

In a similar way the potential with respect to the plasma potential at the origin can be calculated for the boom B extending from $-d_1$ to $-d_2$.

If exact symmetrical conditions are supposed for the two probes, *i.e.* $T_{eA} = T_{eB} = T_e$, $j_{e0A} = j_{e0B}$, $J_{IA} = J_{IB}$ and $J_{pHA} = J_{pHB}$, the following expression for the potential difference between the probes appears

$$\varphi_A - \varphi_B = -\frac{\kappa T_e}{e} \ln \left\{ \frac{\exp \left[\frac{-e}{\kappa T_e} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_1 \right] - \exp \left[\frac{-e}{\kappa T_e} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_2 \right]}{\exp \left[\frac{e}{\kappa T_e} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_1 \right] - \exp \left[\frac{e}{\kappa T_e} (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}_2 \right]} \right\}. \quad (5.11)$$

By series expansions and algebraic manipulations this can be further reduced to

$$\begin{aligned} \varphi_A - \varphi_B &= (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot (\mathbf{d}_1 + \mathbf{d}_2) \\ &= (\mathbf{E} + \mathbf{v} \times \mathbf{B}) \cdot \mathbf{d}. \end{aligned} \quad (5.12)$$

where \mathbf{d} is the vectorial separation between the electrical midpoints of the probes.

It should be pointed out that these calculations should only be regarded as guiding. Nevertheless, this rough analytical treatment is accurate enough to give an understanding of the probe technique, its principles and pitfalls, and it gives a basis for considering the sources of error attached to this kind of measurements.

First of all it should be stressed that the expression for potential difference in Eq. (5.12) appeared under the assumption of exact symmetry in boom orientation and construction. Furthermore, it can be seen that the probe system does not, in general, measure the electric field of an AC wave. It detects only a combination of the field (*i.e.* the gradients of electric potentials) and of the grad-

ients of the various plasma properties which fluctuate under the influence of the wave and thus give rise to variations in the contact potential of the probe. To distinguish between these two contributions is in most cases difficult, as it is necessary to know the complete structure of the wave, and the way the plasma and the probes react to it.

5.5.3. Error analysis. Since the theory for probe measurements of electric fields closely follows the theory for Langmuir probes, most of the pitfalls present in such measurements (*cf. e.g.* Willmore 1970) will also appear in this context. However, as the aims of the experiments are different, the importance of the various errors will be different.

In general, one can say that, since the derivation of the expression for the potential difference between the probes (Eq. (5.12)) depended on the assumption that the probes and their environments could be regarded as exactly alike, the measurements are very sensitive to any effects which will cause asymmetries in the probe configuration. Besides, the presence of the spacecraft and probes will disturb the plasma, and one must consider to what extent this will influence the measurements. Of potential importance for the results will be

- i) The $\mathbf{v} \times \mathbf{B}$ term.
- ii) Disturbances due to finite voltmeter current.
- iii) Effects from spacecraft wake.
- iv) Work function errors.
- v) Sheath overlap.
- vi) Spacecraft potential and interference.
- vii) Asymmetrical photon-emission, shadowing.
- viii) Thermal and aerodynamical bending.

- ix) Contact potentials; errors due to variations in ambient plasma; temperature gradients and particle fluxes.

These various effects and their contributions to the measured voltage will be discussed in Appendix I. It turns out that with careful construction of the probes and with a reasonable dimension of the baseline, it is possible to almost eliminate most of the errors, and to keep the most serious ones down on an acceptable level, < 10% of the total field. However, for vectorial measurements of the field, the $v \times B$ term will place requirements on the determination of the rocket attitude which may be difficult to fulfill.

As a concluding remark it can be said that the probe theory in itself leads to equations and expressions with limited region of validity and great uncertainties. In spite of this it is possible, with the double floating probe technique, to eliminate most of the uncertainties as long as a symmetrical probe configuration can be maintained, and the method has proved to give dependable results (cf. e.g. Maynard 1972, and references given there).

5.5.4. *Experimental set-up for rocket-borne electric field measurements.* This section will give a description of the instrumentation used for the DC and AC electric field measurements performed in the present study. A modified Mini-STEM anten-

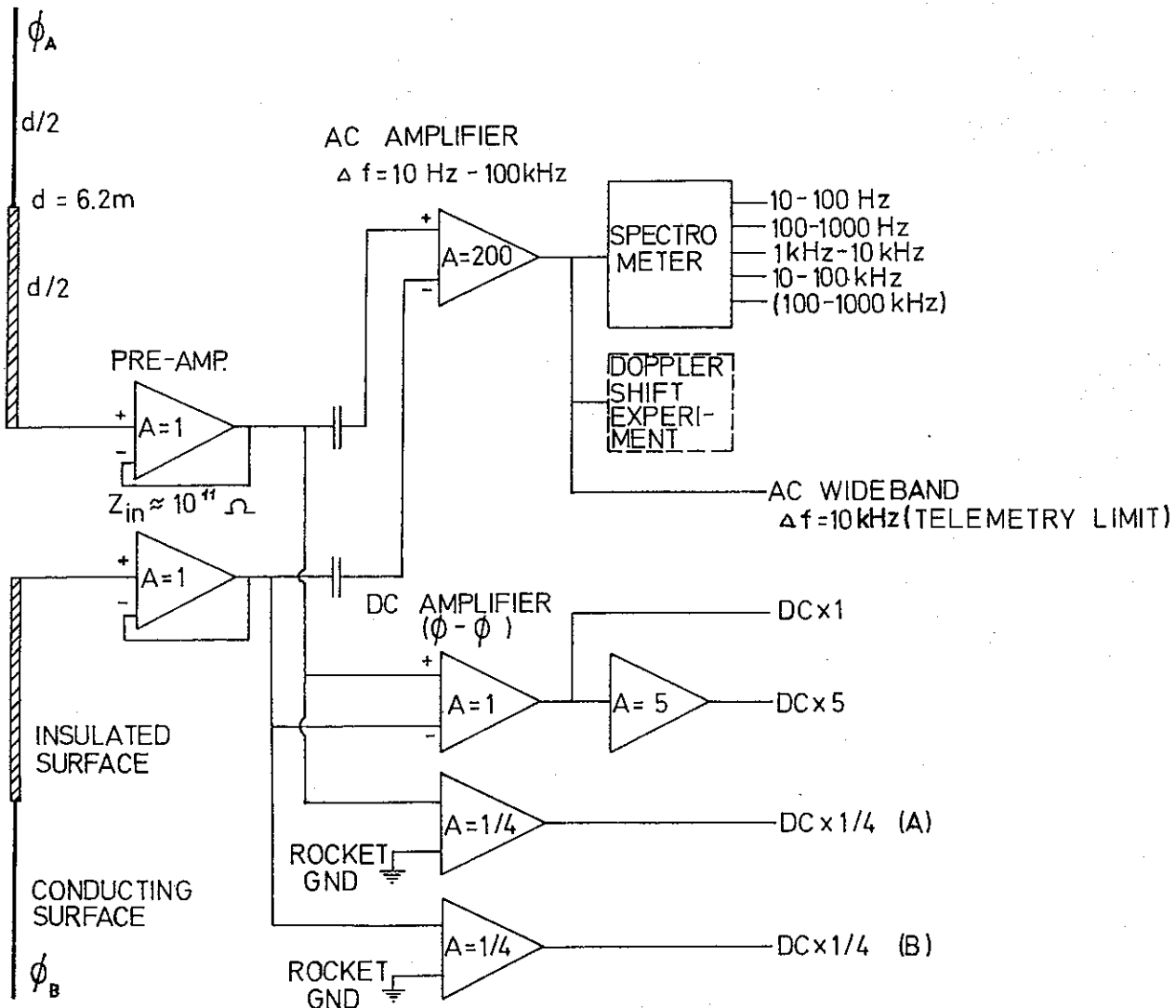


Fig. 5.10. Block diagram of the electronics of the electric field experiment flown in the rockets F23, F24, NASA 14.383 and NASA 18.57. (In the two last rockets the 'Doppler shift' and 'Wideband Channel' were not included.)

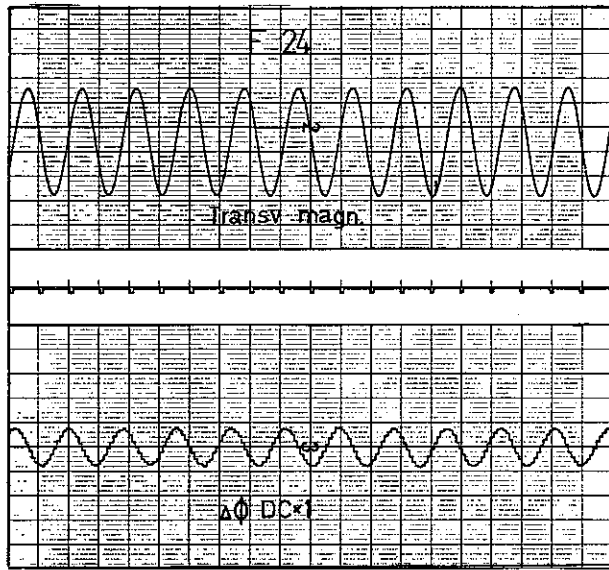


Fig. 5.11. Example of signal from the DC E -field experiment. The sinusoidal form of the record appears from the rotation of the probes, which results in a shift in polarization. The upper curve shows the signal from the rocket's transverse magnetometer.

na has been utilized. In this version the deployment of both antenna elements is controlled by one and the same motor, which assures an even extension for the two booms. The elements, made from beryllium-copper, have noble metal plating to obtain an even work function. The inner half of each element has a special coating, providing DC insulation between the conducting element and the plasma. The boom diameter is 0.007 m and the length 6.2 m, giving a $d_{DC} = 9.1$ m.

The electronics are shown in block diagram form in Fig. 5.10. Precision voltage followers with very high input impedance monitor the antenna voltages. The output of these followers are, in the DC part, subtracted through a precision unity gain differential amplifier giving the desired signal, $\Delta\phi = \phi_A - \phi_B$. By an additional amplifier $5\Delta\phi$ is also monitored. The AC response of the DC electronics goes up to 1 kHz, but in most cases this will be limited by the telemetry capacity.

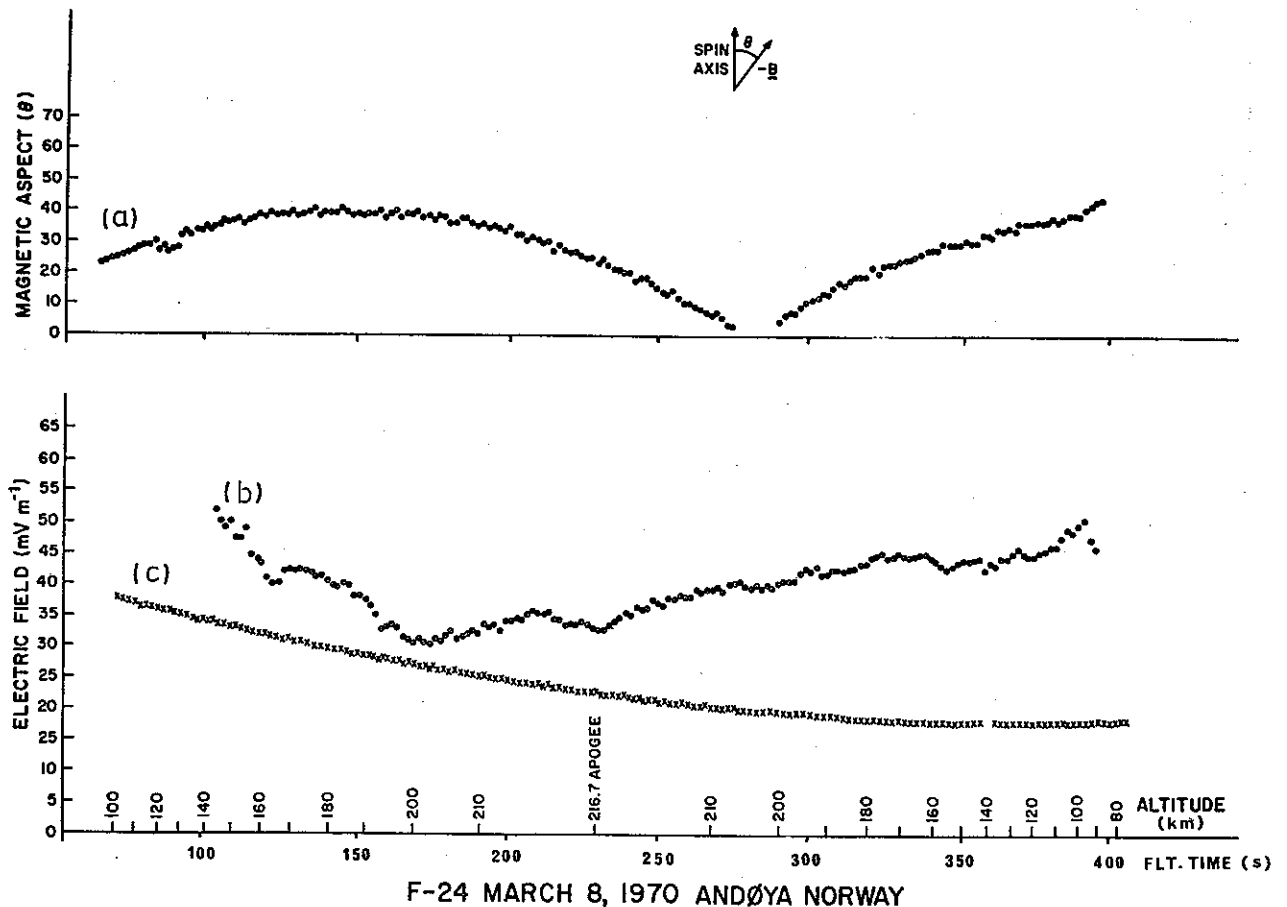


Fig. 5.12. The angle between the rocket spin axis and the negative B is shown in curve (a). Curve (b) shows the total measured field $E + v \times B$, while the calculated $|v \times B|$ component is plotted in curve (c). Data from F24.

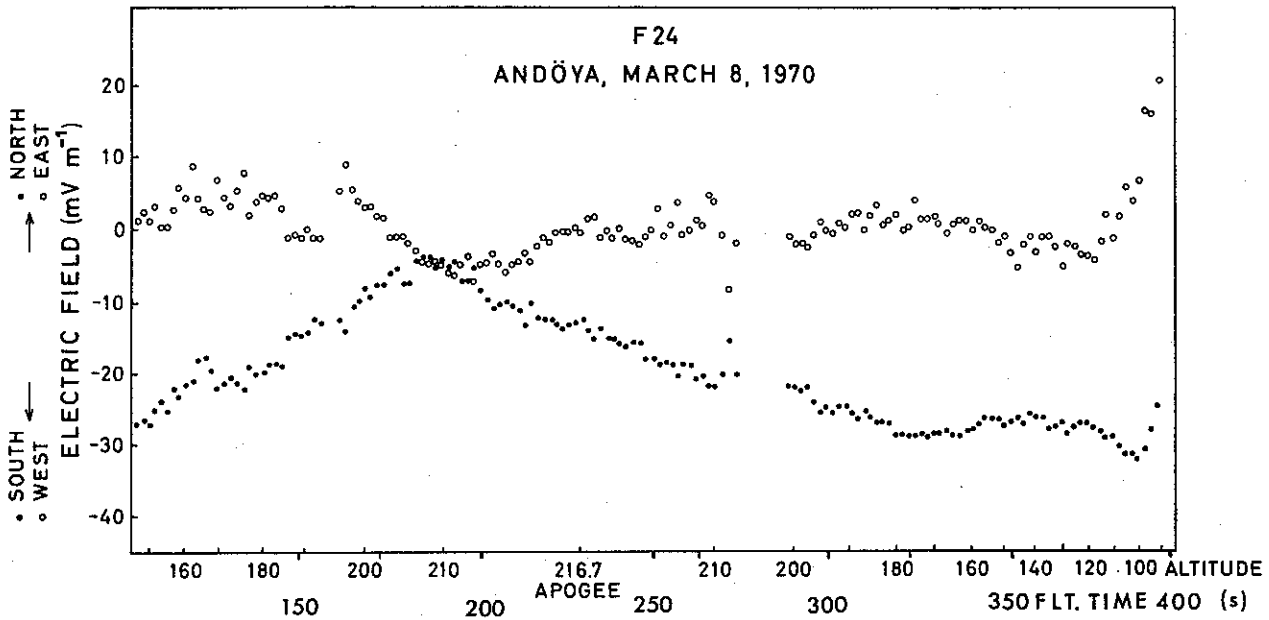


Fig. 5.13. North-south (●) and east-west (○) components of the DC electric field measured during the F24 event. The drop-out in data points between 275 and 290 s is from the part of the flight when the rocket axis was about parallel to the earth's magnetic field, so that determination of the attitude of the probes in a plane perpendicular to \mathbf{B} was very uncertain. The apparent change in the field-components in the last data points is due to bending of the antenna elements caused by atmospheric drag (*cf.* Appendix I).

In addition, the voltage differences between each of the probes and spacecraft ground are also monitored, giving a check on the probe symmetry and variations in contact potentials. The input resistance of the system is greater than 10^{11} ohms and input current leakage less than 10^{-11} A. The capacitance is approximately 50 pF. Comparing this with the findings in Appendix I, one will see that no disturbances from the voltmeter current should be expected. A broadband amplifier with frequency response 10 Hz to 100 kHz is AC coupled to the voltage followers. The AC differential voltages can be monitored directly in broadband analogue form. The AC signal is also recorded by a spectrum analyzer. The number of channels and frequencies covered are subjected to alterations, but in the present version the spectrometer consists of five channels, each covering one decade in the frequency range 10 Hz to 1 MHz. The integrators have a nonlinear response, and the dynamic range is somewhat less than 40 dB.

The sensitivity is of the order of 1 mV m^{-1} for the DC experiment, and $10 \mu\text{V m}^{-1}$ peak to peak for the spectrometer and broadband experiment. The part in the block diagram called 'VLF Dopp-

ler' is a wave propagation experiment which will be further described in Chapter 6.

In the data reduction the magnitude and phase of the total measured DC fields ($\mathbf{E} + \mathbf{v} \times \mathbf{B}$) in the rocket spin plane were determined from the sinusoidally varying voltage from the differential amplifier (see Fig. 5.11) and the lateral aspect magnetometer. To get $\mathbf{v} \times \mathbf{B}$ it is necessary to know the trajectory, which was calculated from slant range Doppler data. The value of \mathbf{B} was taken from the Cain & Cain (1968) model. Vector subtraction of $\mathbf{v} \times \mathbf{B}$ and location of the field in a topographic coordinate system require detailed information on the attitude of the vehicle, which is obtained from the magnetic aspect data.

The errors in the vectorial computations can be large, especially concerning a component that is small compared to the major component (*i.e.* the east-west field in our case). A five-degree attitude error can result in an error of 9% of the spin plane $\mathbf{v} \times \mathbf{B}$. Similarly, the error in the measurement of the phase of $(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ is ± 1 degree, adding another 2% error. Trajectory errors are estimated as $\pm 2^\circ$ in phase $\mathbf{v} \times \mathbf{B}$ calculation or another 4% on the east-west component (making

a total of 15%). The corresponding error from all sources on the main component is +1%. The $\mathbf{v} \times \mathbf{B}$ magnitude early in these flights is about 30 mV m^{-1} . Thus, the confidence level in the results after calculation for the topographic components is $\pm 5 \text{ mV m}^{-1}$. An additional overall error of ± 1 to 2 mV m^{-1} in total magnitude may be present from considerations of the physical technique (*cf.* Appendix I).

5.6. Event studies of DC electric fields

While Sect. 5.4 presented the main characteristics from a great number of electric field measurements, we will in this section examine in more detail the observations from two specific events, with special emphasis on the relation between auroral particles and variations in the field amplitude. The data originate from the two rocket flights (F23 and F24) described in Chapter 2.

5.6.1. Electric fields and optical emissions in the F24 event. Figure 5.12 shows the total measured

field in the spin plane ($\mathbf{E} + \mathbf{v} \times \mathbf{B}$) throughout the flight. In Fig. 5.13 the calculated east-west and north-south components of the field are plotted. The measurements seem to reflect the stable PCA launch conditions, with almost total lack of short period variations. The field amplitude varied from a few mV m^{-1} to near 40 mV m^{-1} , with the lowest values recorded from 150 to 250 s (flight time). No systematic altitude dependence was observed. The field was, within the accuracy of the attitude determination, southwards, consistent with the established picture of a westward Hall current during a negative bay.

The optical instrumentation in the payload consisted of three photometers, mounted with their optical axis parallel to the rocket's longitudinal axis. The emissions $\text{N}_2^+ \text{IN}$ (4278 \AA), OI (5577 \AA) and $\text{N}_2 \text{IP}$ (6705 \AA) were measured. The measured photon emission rate above the rocket at 4278 \AA during the flight is shown in Fig. 5.14. The observed asymmetry of the curve about apogee cannot be explained by temporal variations of

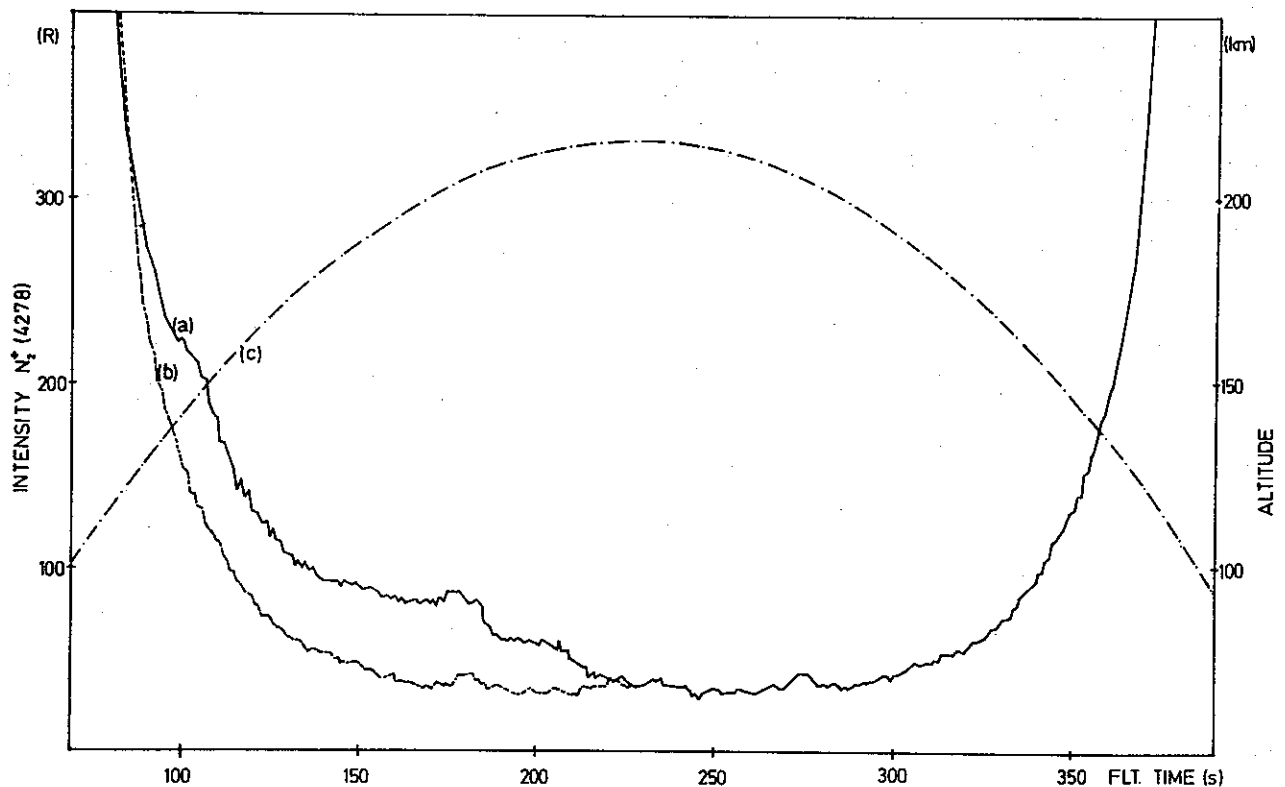
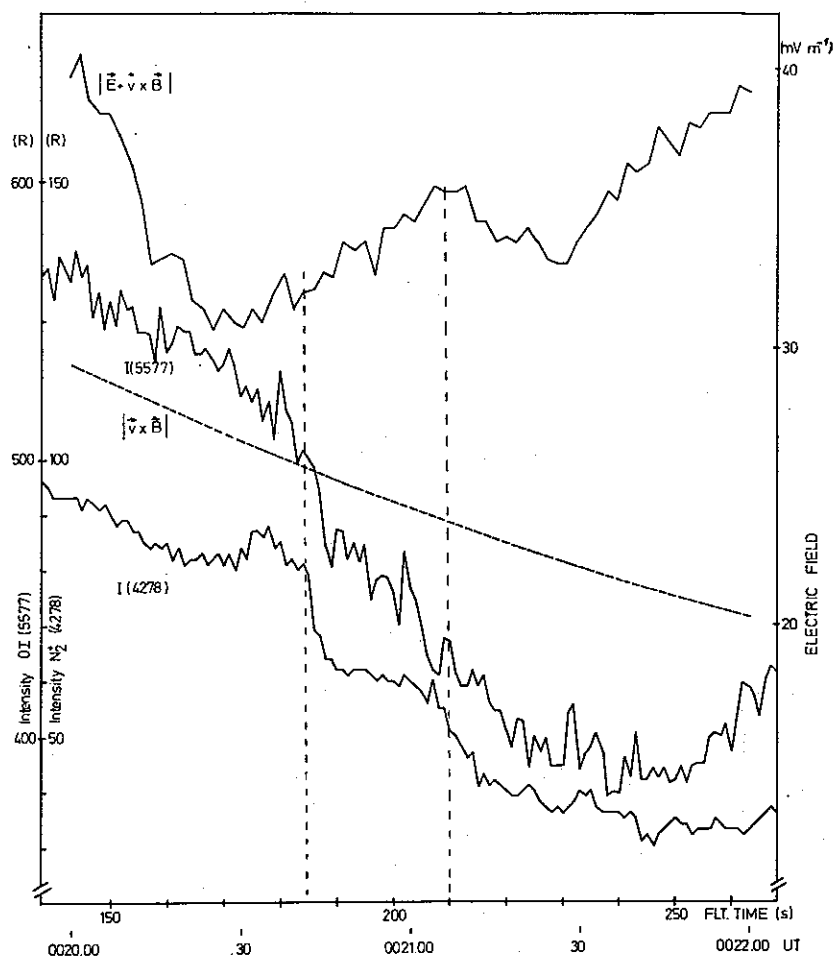


Fig. 5.14. Curve (a) shows photometric recordings of the intensity of N_2^+ (4278 \AA) as measured above the rocket *vs.* flight time. In curve (b) the intensity measured during the descent part of the flight is mirrored about apogee (228 s) to show the asymmetry in the recordings. The rocket altitude *vs.* time is plotted in curve (c). Data from F24.

Fig. 5.15. Observed variations in auroral intensities and the electric field when the rocket F24 penetrated the structural region. The field values are averaged over one spin period (1.8 s) while 1 s smoothing is used for the optical emissions. The locations of the auroral forms are indicated by the vertical lines.



the aurora (*cf.* Fig. 2.1.c), or variations in the rocket attitude, and it is therefore concluded that the rocket in the flight period 150 to 220 s penetrated a region where auroral structures were superimposed on the homogeneous glow (*cf.* Måseide *et al.* 1973).

Figure 5.15 shows a high time-resolution plot of the field and photometer recordings at 5577 Å and 4278 Å when the rocket was in the structural region. The electric field is averaged over one spin period (1.8 s) while the optical recordings are smoothed over 1 s. It can be seen that the rocket passed through a double structure in the DC electric field, with two minima at 175 and 230 s separated by a small maximum at 210 s. Notice also, especially between 160 and 200 s, that some small-scale fluctuations in the range 0.1 to 0.3 Hz are observed.

The main features in the optical recordings are:

- i) During the two intervals, 150–170 s and 190–210 s, the light intensity decreases fairly steadily.
- ii) Two abrupt changes at 185 and 210 s, respectively, are seen in the 4278 Å emission, while the 5577 Å intensity decreases markedly slower.
- iii) A small maximum is found between 175 and 185 s in the N_2^+ emission which is not observed in the oxygen line.

The assumed auroral situation can be represented in a somewhat simplified picture, qualitatively shown in the geometrical sketch in Fig. 5.16. Two auroral structures of roughly 5 km widths and separated by about 20 km are penetrated by the rocket. This picture explains the 4278 Å observations fairly well. Comparisons

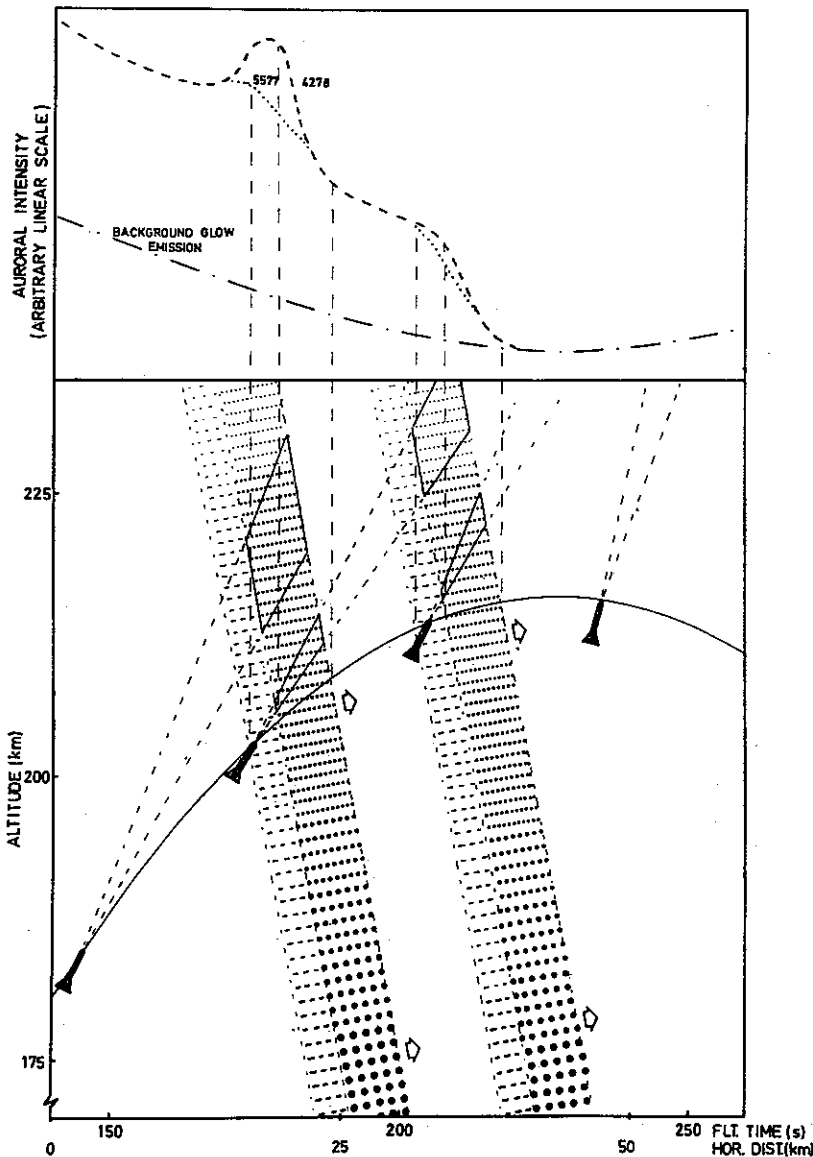


Fig. 5.16. A qualitative sketch for the F24 event, showing the assumed geometrical configuration between the auroral structures and the photometer's field of view when the rocket penetrates the auroral forms. The corresponding optical recordings are shown at the top of figure in a simplified version. The auroral structures are represented by dashed (5577 Å) and dotted (4278 Å) areas, indicating a partly spatial separation of the OI (5577 Å) and the N_2^+ (4278 Å) emitting regions due to auroral movement.

with the green oxygen line indicate that the auroral structures move in the same direction as the rocket (*i.e.* northward motion). This conclusion is based on the fact that the lifetime for the excited OI (1S) is about 0.7 s compared to a lifetime of the N_2^+ ($B^2\Sigma_u^+$) being of the order of 10^{-7} s (*cf. e.g.* Omholt 1971); thus the OI (5577) emission has a natural time lag compared to the N_2^+ (4278) emission. This time lag together with a northward movement of the auroral forms can explain the difference observed between the two emissions.

From Figs. 5.15 and 5.16 the following conclusions concerning the DC electric field in and near the auroral forms can be drawn.

- The field decreases as the rocket approaches the auroral structure and it reaches a minimum of just a few $mV m^{-1}$ when the rocket enters the first auroral form. The field minimum is actually observed when the rocket is in the 5577 Å time lag structure.
- Between the two forms, the field increases slightly, and a peak is observed when the rocket enters the second auroral form.
- Again the field starts to decrease and a minimum is observed when the rocket passes the northern edge of the second form.
- During the rest of the flight the field increases slowly while the overall auroral emissions decrease correspondingly.

In comparing these results with previous findings of E -field behaviour in auroral forms (*cf.* Sect. 5.4) the overall glow aurora must be considered in addition to the weak enhancements. Any model using loading of a current generator depends on enhanced ionization in the E -region, since the perpendicular conductivities in the F -region will be too low, due to reduced collision frequencies, to permit a significant increase in the current. The increase in E -field and the reduced auroral intensity during the last part of the flight indicate this effect. The observed relation between the aurora and the E -field is, however, somewhat confused, with a minimum of the field inside the first form and a local maximum inside the second one. Even though the rocket was above the E -region when penetrating the auroral forms, the particle spectrum in these forms will most probably also contain particles with sufficient energy to penetrate to lower altitudes and cause E -region ionization. The particle flux and the extent of the ionization can, however, be different in the two forms. A more unified picture may also be obtained if the whole region, from the southernmost boundary of the first form to the northernmost rim of the second one, is defined as *one* 'auroral region'. In this case the overall behaviour of the field will be that it increases with decreasing glow aurora.

However, in the auroral forms superimposed on the glow, the field has local minima at the edges of the structured region. The direction of the field is, in the main, preserved inside the aurora, but a slight shift towards southwest can be seen.

5.6.2. *Electric fields and particle measurements in the F23 event.* The electric field observed in the F23 event is, in accordance with the general ionospheric conditions, much more variable over short distances and/or time. Figure 5.17 shows the total field measured in the spin plane, together with the magnitude of the $\mathbf{v} \times \mathbf{B}$ component. The conspicuous change in field strength which takes place around 2223.45 UT naturally divides the record in two parts, a low intensity region before this change and a more vigorous last section of the flight. In the first part variations in the field from a few mV m^{-1} to approximately 20 mV m^{-1} takes place, and the general trend in the record is towards a steadily increasing field.

With the sudden change in field strength the amplitude reaches a magnitude of almost 100 mV m^{-1} . From this peak value a decrease of the field, with rapid fluctuations with peak to peak amplitude of about 5 mV m^{-1} superimposed, takes place. Towards the end of the flight some more violent field variations again appear.

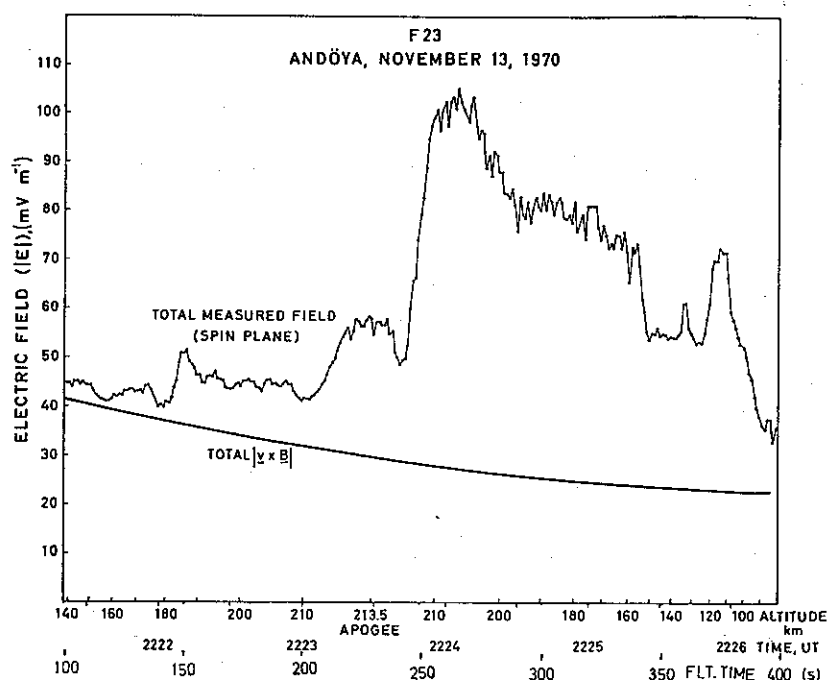


Fig. 5.17. Total DC electric field measured in the spin plane for the rocket F23. The calculated $|\mathbf{v} \times \mathbf{B}|$ contribution is also shown.

The drastic change in electric field strength happened approximately at the same time as the sudden increase in magnetic H -deviation and riometer absorption started. According to the on-board low energy particle detectors this event did not give rise to an increased precipitation of low energy (<15 keV) electrons (*cf.* Maynard *et al.* 1973). The ground observed magnetic field variations are therefore in this case believed to be more associated with an increased electric field, than with changes in the ionospheric conductivities.

Before the abrupt change in the field strength the general tendency in the record was for the field to slowly increase. The detectors monitoring low energy particles (protons and electrons in the range $\sim 0.5 - \sim 15$ keV) showed at the same time on the average a steady decrease in the countrate (Maynard *et al.* 1973). Superimposed on this declining curve are, however, several short time variations, most having no obvious relation to the changes in the electric field. Of interest is, however, a peak in the electron spectrum between approximately 3 and 6 keV which appears between 150 and 200 s in flight time (Fig. 5.18). If we compare this with the E -field record we find that around these two times small minima appear in the field strength and that the amplitude stays at a somewhat increased level in the interval between. As electrons in just this energy range are most actively involved in the production of oval aurora (Deehr *et al.* 1970), this region may well be associated with an auroral structure. If this is the case a situation quite similar to that which appeared in the F24 event is present, *i.e.* the field has a minimum value near the boundaries of an auroral region.

5.6.3. Concluding remarks. The observations do not bring forward any clear relation between the electric fields, aurora, and auroral particles. In the main they seem to support the assumption of a decreased field in auroras, but this picture can be rather mixed up. One should, however, be aware of the importance of the height of the auroral emissions, or the energy of the primary particles, in this connection. Local effects, such as polarization fields and potential differences caused by different parallel conductivities along

different L -shells must also be taken into account. Plasma instabilities may also give rise to perturbations in the large-scale electric field (*cf.* Sect. 8.2).

An interesting feature in the observations is the apparent minima seen in the field at the edges of an auroral region. Although small, these seem to be significant. In the current model proposed by Park & Cloutier (1971) (*cf.* Fig. 5.2), antiparallel Birkeland currents flow at the sides of an auroral region. It thus appears that the observed minima in the field may coincide with the location of Birkeland current regions. To make further con-

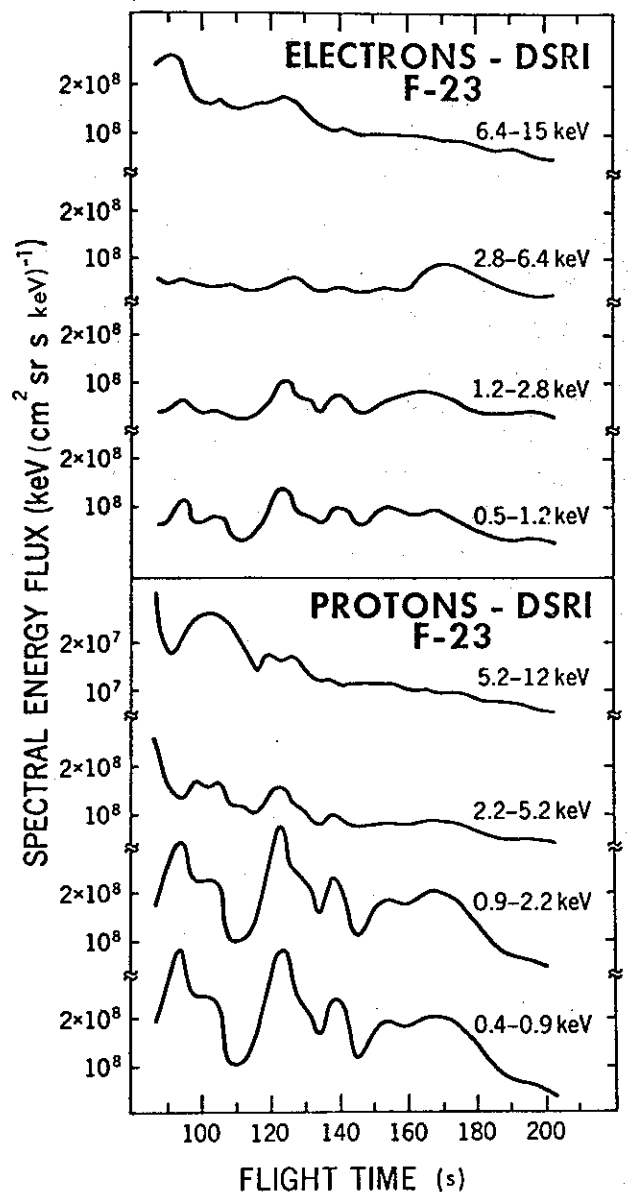


Fig. 5.18. Spectral energy fluxes as function of flight time for various electron and proton energies, measured during the flight of F23 (after Maynard *et al.* 1973).