

Space Physics

Carl-Gunne Fälthammar

Second Edition, Third printing
Stockholm August 2001

SPACE PHYSICS

Contents

Preface	iii
1. Introductory Survey	1
1.1 The Role of Space Observations	1
1.2 The New Concept of Space	2
1.3 The Plasma State	6
1.4 Regions of Space	11
2. The Atmosphere	13
2.1 Origin	13
2.2 The Standard Atmosphere	13
2.3 Scale Height	17
2.4 Structure of the Atmosphere	19
2.5 Solar Radiation	21
2.6 Electric Conductivity	23
3. The Ionosphere	24
3.1 History	24
3.2 The Ionospheric Layers	27
3.3 Ionospheric Disturbances	32
3.4 Radio Wave Propagation in the Ionosphere	34
3.5 Radiative Ionisation in an Exponential Atmosphere	38
3.6 Diffusive Equilibrium in Multicomponent Plasma	40
3.7 Motion of Charged Particles in Electric and Magnetic Fields	43
3.8 Electrical Conductivity in Magnetized Plasma	48
3.9 The Appleton–Hartree–Equation	53
4. The Magnetosphere	60
4.1 The Geomagnetic Field	60
4.2 The Structure of the Magnetosphere	64
4.3 The Earth's Radiation Belts	67
4.4 Dynamics of the Magnetosphere	70
4.5 The Aurora	74
4.6 Mathematical Models of the Geomagnetic Field	84
4.7 Field Transformations	88
4.8 Motion of Magnetically Trapped Particles	91
4.9 Other Magnetospheres	99
5. The Sun	110
6. Interplanetary Space	122
6.1 The Solar Wind	122
6.2 Gas Dynamic Solar Wind Analogy	129
6.3 Frozen-in Magnetic Field Lines	131
6.4 The Critical Velocity	133

7. Stellar and Interstellar Plasma	137
8. Intergalactic Plasma	142
9. Cosmic Radiation	144
10. Space as a Laboratory	149
Bibliography	154
Handbooks	154
Monographs	154
Popular Publications	155
Scientific Papers	157
Subject Index	164
Author Index	167
Appendix	168
Natural Constants	168
Common Units	168
Index of Notations	169

Preface

The present text is based on a general introductory course for undergraduate students, given annually since 1989 at the Royal Institute of Technology (KTH).

Following a strong student interest in space-related science and technology the course was developed at the Department of Plasma Physics of the Alfvén Laboratory, KTH. It constitutes the first of a group of courses arranged for students wishing to specialize in space-related subjects.

Stockholm, August 2001

Carl-Gunne Fälthammar

1. Introductory Survey

1.1. The Role of Space Observations

Since rocket technology made it possible to place scientific instruments in space, our possibilities of observation have expanded in two essential respects:

(1) New Spectral Ranges

Information about distant celestial bodies used to reach us only within two very limited wavelength bands of the electromagnetic spectrum, namely the visual band, about 380 – 780 nm (3800 – 7800 Å), and the high frequency end of the shortwave radio frequency range. Only within these “windows” can the radiation pass through the atmosphere and be observed at the Earth’s surface, see Fig. 1.1.1. All information carried by infrared, ultraviolet, X-ray and gamma-ray waves was excluded.

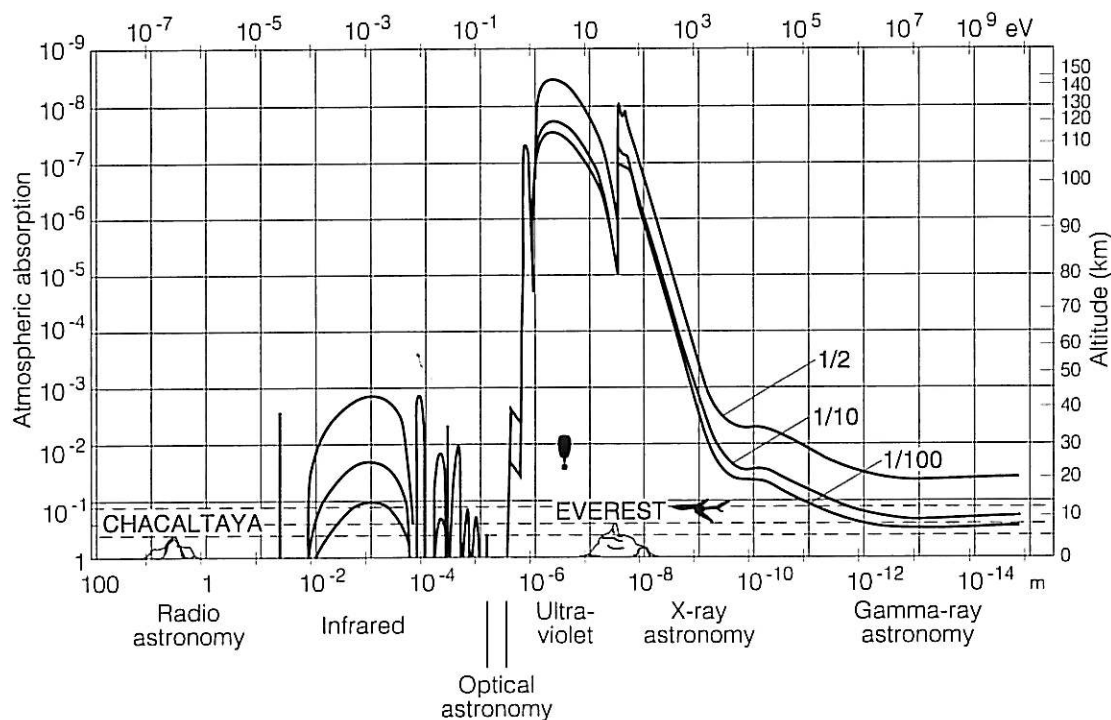


Figure 1.1.1. Atmospheric absorption in different wavelength ranges. The curve shows the altitude above sea level to which the radiation can penetrate.

(2) Measurements *in situ*

Direct measurements in space have made it possible to acquire knowledge of physical conditions that cannot be detected remotely. For example, electric and magnetic fields, as well as the extremely thin, and therefore invisible, matter, the space plasma, can only be observed by means of direct measurements on the spot, *in situ*. It is only thanks to such measurements that we have learned about these quantities, their structure and other properties, and about the complicated physical processes that take place in the space plasma.

The new information obtained from space age observations has drastically changed our concepts of the near-Earth space environment as well as of the universe as a whole. This introductory chapter will give a brief survey of these new concepts.

1.2. The New Concept of Space

Before the space age our concept of outer space was a very simple one. Space was assumed to be essentially a vacuum, whose content of matter was limited to the extremely high energy particles that constitute the cosmic radiation (Chapter 9). (The cosmic radiation represents a material density of only about $3 \cdot 10^{-17}$ particles per m^3 .)

In conformity with this view it was further believed that (other than the radiation fields of stellar light and cosmic radio noise) the only electromagnetic fields in space were the magnetostatic fields that originate from electric currents in the interiors of the Earth and other celestial bodies. According to a well-known theorem in electromagnetism the magnetic field from a bounded current distribution becomes more and more dipole-like with increasing distance from the source. The Earth's magnetic field is very nearly a dipole field at the Earth's surface. It was assumed to become increasingly dipole-like with increasing distance from the Earth (Fig. 1.2.1a). As the strength of a dipole field decreases as the inverse third power of radial distance, it should soon become negligibly small.

One of the great pioneers of space research, Carl Störmer, analysed the orbits of electrically charged particles in the geomagnetic dipole field (Störmer, 1913, 1955, see *e.g.* Alfvén and Fälthammar 1963, pp. 43–46). His intention was to explain the aurora ("Northern lights"), but the results are mainly applicable to orbits of cosmic rays.

One of Störmer's findings was that in a magnetic dipole field there exist so-called *forbidden regions*. In such a region an electrically charged particle can remain trapped indefinitely. By reversing time in the equations of motion of the particle, it can be shown that there is also no way the particle can get *into* the forbidden region from outside. It was therefore taken for granted that also in the real geomagnetic field there were no such trapped particles.

The first surprise encountered when the satellite Explorer 1 was sent into space was that Störmer's forbidden regions - instead of being empty - contained such a

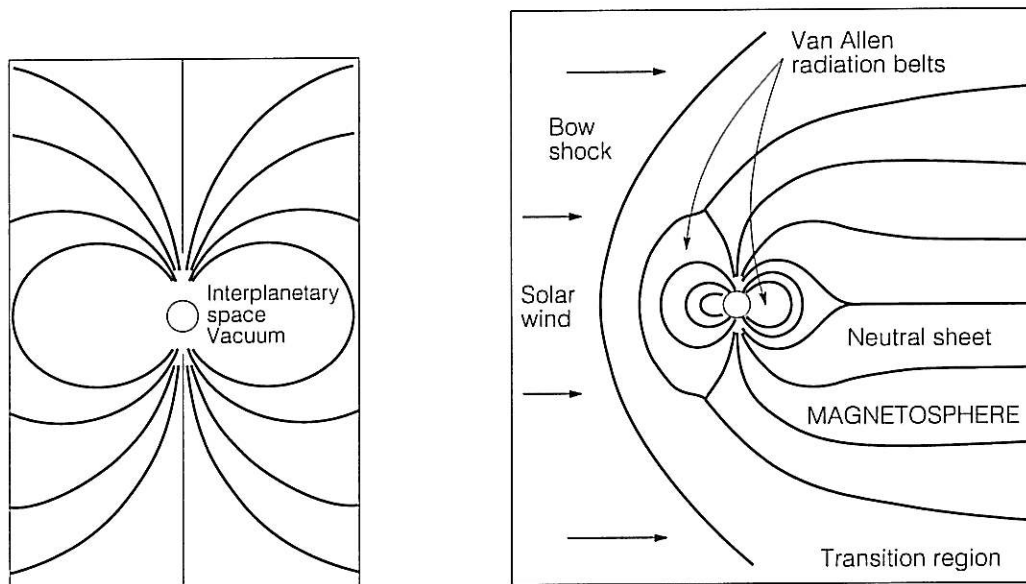


Figure 1.2.1a. Space in the 1950's: A structureless vacuum threaded by a rapidly diminishing magnetic field.

Figure 1.2.1b. Space in the 1960's: A deformed geomagnetic field with radiation belts (Van Allen belts) of high energy magnetically trapped particles.

profusion of high energy trapped particles that they constitute what we now know as the *Earth's radiation belts* (*Van Allen-belts*). The radiation was so intense that the satellite's Geiger counter was saturated and therefore gave zero reading at high altitudes.

As for the Earth's magnetic field, it was found that, in accordance with the text book, it became increasingly dipole-like with increasing distance from the Earth. But at a distance of about 10 Earth radii from the Earth there was a new surprise: the magnetic field changed abruptly, often to the opposite direction. This "discontinuity" (not always equally pronounced), which was long referred to as the *Cahill discontinuity* after its discoverer, is, as we now know, the boundary between the Earth's magnetic field and an external, interplanetary, magnetic field. Continued mapping of the Earth's magnetic field by means of satellite-borne magnetometers showed that it was severely deformed and confined to a cavity in space, as shown in Fig. 1.2.1 b. This cavity, to which the dominating influence of the Earth's magnetic field is confined, is now called the *magnetosphere*.

The discovery of the radiation belts was followed by extensive exploration of the occurrence and properties of high energy trapped particles. The original two Van Allen Belts were followed by a multitude of trapped particle populations, which were found to fill large parts of the magnetosphere. But for a long time (largely for lack of suitable detector technology) no measurements were made of what we now know to be the most important component of matter in the magnetosphere, namely *plasma* consisting of particles with energy much lower than that of radiation belt particles. These plasma particles have energies in the keV range or below, corresponding to an *equivalent temperature* of 10^7 K or less. (Equivalent temperature can be defined by $(3/2)kT_{eq} = W = eV$, where W is the particles' average energy in Joule and V its voltage equivalent, k is Boltzmann's constant, $k = 1.38 \cdot 10^{-23}$ J/K, and e is the electron charge, $e = 1.602 \cdot 10^{-19}$ C. The ratio e/k is 11609 K/volt.)

With increasingly sophisticated measurements more and more of the complexity of the magnetosphere emerged. We now know that the magnetosphere is richly structured (Fig. 1.2.1c) and contains a multitude of plasma populations, whose *temperatures* cover more than six powers of ten, from about 1000 K in parts of the ionosphere (Chapter 3) to (equivalent) temperatures more than 10^{10} K (particle energies of several MeV) in parts of the radiation belts. Their *densities* also cover more than six powers of ten. Thus, the number of particles per unit volume varies from more than 10^{12} m^{-3} in the F-layer of the ionosphere to 10^4 or less in parts of the polar plumes of the magnetosphere.

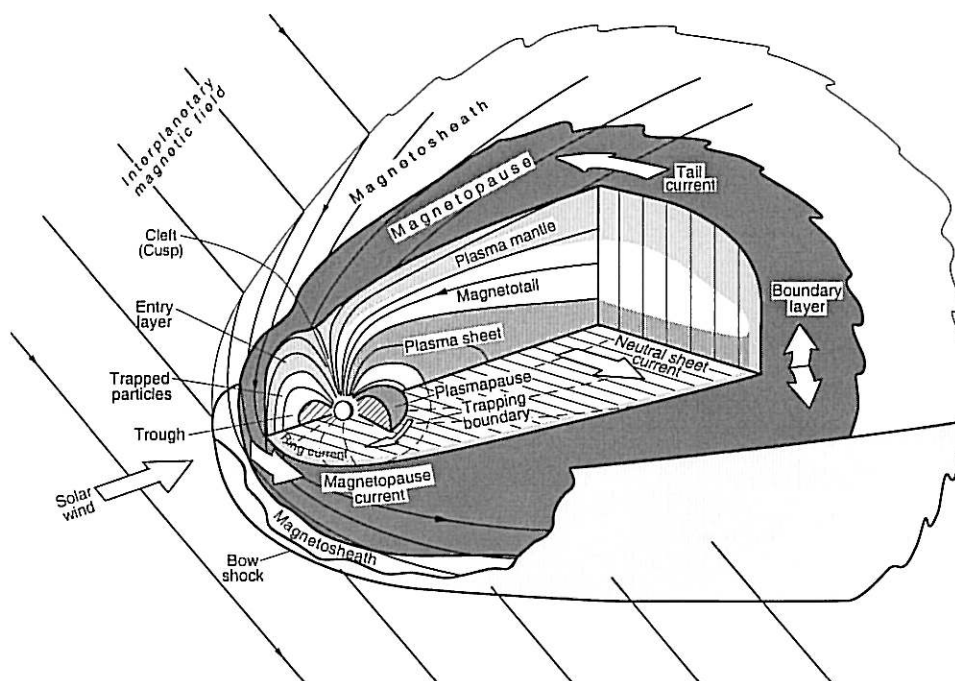


Figure 1.2.1c. Present-day space: A richly structured magnetosphere, containing a multitude of plasma populations, in which complicated plasma processes take place. The manifestations of these processes range from natural phenomena such as the aurora (Northern lights) to consequences such as damage to communications satellites and disturbances in Earth-bound electrical systems.

But even after the “modern” concept of the magnetosphere was well established (including many of the plasma regions shown in Fig. 1.2.1c), there still prevailed a serious misconception about matter in space. It was taken for granted, that the ions observed in the magnetosphere were hydrogen nuclei (protons), *i.e.* that the plasma filling the magnetosphere was a hydrogen plasma from the sun, carried to the vicinity of the Earth as a solar wind (Chapter 6). Only rather recently, after the chemical composition was also measured, it was realized that large parts of the magnetosphere is sometimes dominated by oxygen plasma expelled from the ionosphere (with the Earth’s own atmosphere as the ultimate source). This discovery implies, among other things, that in the space plasma around the Earth, there exist efficient but still incompletely understood mechanisms of chemical separation, probably driven

by electric fields and/or plasma waves. The existence of these mechanisms can have far-reaching consequences also for the understanding of plasmas in remote parts of the universe, which can never be reached for direct measurements.

One reason why the Earth's magnetosphere is so interesting from a physics point of view, is the abovementioned rich variety of plasma populations. But even more important is the fact that in these plasma populations there occur a multitude of plasma physical processes that are of fundamental importance for the understanding of the behaviour of matter in the plasma state. The reason for this is, in its turn, that the relatively cool and dense plasma of the ionosphere, which constitutes the lower boundary of the magnetosphere, is tied by friction to the Earth's upper atmosphere, whereas the outer, hotter parts of the magnetosphere have a strong dynamic coupling to the solar wind plasma that streams past it. The interaction between the outer and inner parts of the magnetosphere takes place primarily by electric currents (a few million Ampères) along the geomagnetic field lines. These currents provide exchange of both energy and momentum. But in addition they cause, both directly (by the exchange of charge carriers) and indirectly (as a consequence of instabilities), an exchange of matter (plasma), which is highly selective and leads to the abovementioned chemical separation.

We now know that practically all matter in the universe is in the plasma state, which can be considered the fourth state of matter, as described in section 1.3. The properties of matter in the plasma state are still unknown in important respects. This is particularly true for the kinds of plasma that are present in space. The background is the following.

On the basis of experiments in "cool" and weakly ionized plasmas – the only ones that could be produced at that time – the so-called classical plasma theory had been developed already in the 1950's. It was a generalization of the ordinary kinetic theory of gases, where the electric and magnetic forces on the charged particles had been included. Although, for technical reasons, it had not been possible to test the theory at very high temperatures, it was assumed to have general validity. It was on the basis of this theory that, in the 1950's, it was predicted that energy production through *thermonuclear fusion* would be a reality within 15 years. Thermonuclear fusion requires that the fuel has a temperature of about a hundred million K, and great resources were created to produce and confine extremely hot plasma. As soon as experiments with hot plasma became possible, it was found that it did not behave as expected but exhibited one "anomalous" phenomenon after the other. It became necessary to start again from the beginning, with theory and experiment progressing hand in hand.

At about the same time it became possible to study space plasmas by means of instruments carried on rockets and satellites. Fusion research and space research together opened the era of what we now call modern plasma physics. This branch of science is still in a stage of intense basic research, with many fundamental problems still unsolved.

The development of the new concept of the universe, which has been acquired through space research, can be characterised as a transition from an "Optical Universe" (derived from information carried by visible light) to a "*Plasma Universe*" (which emerges when the whole electromagnetic spectrum is used). A popular description of this "change of paradigm" has been given by Alfvén (1986).

A consequence of the new paradigm is that investigations that can be carried out in the neighbourhood of the Earth, are also of fundamental importance to astrophysics, because they provide a more reliable basis for the physical interpretation of astrophysical observations (Fälthammar *et al.* 1978; Alfvén 1981; Fälthammar 1988a,b). As the plasma state must have dominated also during the evolution of the universe to its present state, this has consequences also for *cosmogony* (the science of the origin and evolution of planet and satellite systems) and *cosmology* (the science of the origin and evolution of the universe) (Alfvén 1981; Fälthammar 1989b).

1.3. The Plasma State

The States of Aggregation

Matter can exist in three different *states of aggregation*: the solid, liquid and gaseous states. These are characterised by different average energy per particle (atom or molecule). Solid matter has the lowest energy content, gaseous matter the highest. The orders of magnitude of the energy per particle are illustrated in Fig. 1.3.1 by the example of water. Transition energies are given in electron volts per molecule ($1 \text{ eV} = 1.602 \cdot 10^{-19} \text{ J}$).

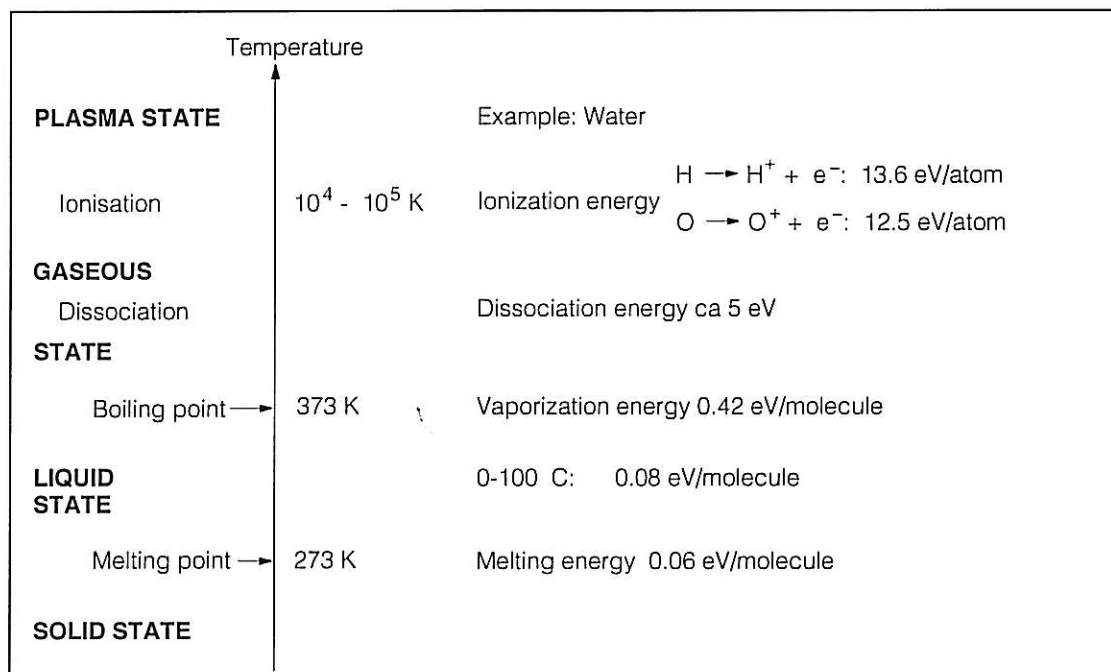


Figure 1.3.1. Comparison of transition temperatures and transition energies for the transitions solid – liquid, liquid – gas and gas – plasma.

If matter in the gaseous state is further heated, molecules will be dissociated, and, when the thermal motion becomes violent enough, more and more atoms will be split into their electrically charged constituents, electrons and positive ions – the gas becomes ionized and transformed to a *plasma*. As shown in Fig. 1.3.1, rather large amounts of energy per particle are then involved. To ionize an oxygen or hydrogen atom takes about 13 eV, which is more than two hundred times the melting energy and more than thirty times the vaporization energy.

Transitions between the three states of aggregation occur at transition temperatures, melting and boiling points, which at a given pressure are very well-defined. In contrast, the temperature required for ionization of a gas is not well-defined. The transition from unionized to ionized state takes place gradually over a large temperature interval. Furthermore, ionization is often due to other causes than heat, for example electron impact or short wave electromagnetic radiation.

Sometimes the plasma state is referred to as a state of aggregation of its own, the *fourth state of matter*. From a strict thermodynamic point of view this terminology is somewhat inappropriate, since the transition between gas and plasma is so different from the genuine phase transitions melting and boiling. On the other hand, the presence of electrically charged particles causes the plasma to have drastically different properties than an unionized gas. From that point of view the plasma does deserve the status of a state of aggregation.

The name *plasma* derives from the greek word *plassein*, which means “to shape”. It was introduced by Irving Langmuir to denote the positive column of a gaseous discharge, where the “plasma” is a luminous region, whose extent adjusts itself to available space. In modern terminology plasma simply means ionized gas (except for the meanings the word has in other branches of science, such as medicine and solid state physics).

Occurrence

In our daily environment the three “ordinary” states of matter dominate, and natural plasma is met with only occasionally. One example of natural plasma is a lightning discharge, where temperature can rise to 25 000 K and lead to an ionization degree of 80 %. Another example is meteor tracks, (the plasma properties of which can be used for communication purposes by reflexion of radio waves).

For the universe as a whole, the opposite is true. Solid, liquid and gaseous matter exists on cold celestial bodies, such as planets, satellites, asteroids, comets, meteoroids and other small bodies, down to microscopic dust grains. But all these celestial bodies together represent only a vanishingly small part of the matter in the universe. Vastly more mass resides in stars on one hand and diffuse interstellar matter on the other, in both cases in the plasma state.

Definition

Since the number of charged particles – both in absolute and relative terms can vary within wide limits, one may ask what is required for matter to count as plasma.

The answer to this question depends not only on the density of charged particles but also on the relevant dimension of the phenomenon concerned. An astronaut in interplanetary space, where there are only 1 – 10 electron-ion pairs per cm^3 , will experience his surroundings as a nearly perfect vacuum. But in their large-scale streaming around the Earth and other planets, these charged particles constitute the solar wind, which behaves like a continuum and can be described in terms of equations similar to those of gas dynamics, although more complicated.

It can be shown that a crucial parameter for determining whether plasma conditions prevail is the *Debye length*, λ_D , which is defined by

$$\lambda_D = \sqrt{\frac{\epsilon_0 k T_e}{n_e e^2}} \approx 69 \sqrt{T_e / n_e} \text{ m} \quad (1.3.1)$$

where n_e is the number of electrons per m^3 and T_e is the electron temperature in K. In phenomena with a characteristic length l_c matter behaves essentially as a plasma if

$$l_c \gg \lambda_D \quad (1.3.2)$$

otherwise not. Sometimes it is also required that the number of charged particles in a sphere with radius λ_D , the *Debye sphere*, is large, *e.g.* $n\lambda_D^3 \gg 1$. This is, however, is usually trivially satisfied.

In the above-mentioned case of the solar wind one finds that λ_D is of the order of magnitude of ten meters. Thus phenomena in the above mentioned astronaut's local surroundings do not satisfy the condition $l_c \gg \lambda_D$, while the large-scale streaming of the solar wind does so with a large margin.

Quasineutrality

In a plasma the positive and negative charges must balance each other very closely, except in very small regions. This property of the plasma is called *quasineutrality*. It can be shown that the relative unbalance $\Delta n/n$ must satisfy the condition

$$\Delta n/n < (\lambda_D/l_c)^2 \quad (1.3.3)$$

The reason is that a greater charge imbalance would create unreasonably large electric fields. Since in a plasma, by definition, $l_c \gg \lambda_D$ quasineutrality implies an extremely close balance between positive and negative charges.

Complexity

The complicated behaviour of matter in the plasma state is based on the motion of the electrically charged particles that constitute its key ingredient.

In an ordinary gas each particle (atom or molecule) moves in ballistic, practically rectilinear, orbits between nearly instantaneous collisions, at which the direction and speed of the motion changes, so that the particle enters another rectilinear orbit. This simple behaviour is a result of the atomic forces having an extremely short range.

In a plasma the orbits of the (charged) particles are instead extremely complicated. This is so for two main reasons:

- 1) The electrostatic force between electrically charged particles decreases only as the square of the distance between them, and therefore has a very long range. This means that every particle at every instant of time interacts electrostatically with many, usually thousands of, other particles.

Fortunately it is possible (see for example Alfvén and Fälthammar 1963 section 4.3) to define an “equivalent” collision cross section, by means of which the mutual interaction between the charged particles can be described as if it were made up of ordinary collisions. It would lead too far to discuss this in the present context, and here will only be given an example: For the interaction of an electron with positive ions with the charge number Z , the *equivalent cross section*, S_{eq} , is given by

$$S_{eq} = 0.8 \cdot 10^{-9} Z^2 \ln \Lambda / T_e^2 \quad (1.3.4)$$

where the so-called *Coulomb logarithm*, $\ln \Lambda$, is of the order of magnitude 10 and varies only slowly with density and temperature.

Comparing the equivalent cross section, S_{eq} , with the collision cross section for collisions between neutral atoms and molecules, which is of the order of 10^{-19}m^2 , one finds that at moderate temperatures (of the order of ten thousand K) the (equivalent) cross section is much larger for charged particles than for neutrals. An important consequence of this is that a plasma can behave as if it were almost fully ionized already at ionization degrees of a few percent or even tenths of percent.

- 2) Even apart from the mutual interaction of the particles, the charged particles have a complicated motion, because they are influenced by the magnetic fields that are present in practically all space plasmas.

A particle with charge q and velocity \mathbf{v} perpendicular to a magnetic field \mathbf{B} is subject to the force $q\mathbf{v} \times \mathbf{B}$, which has the magnitude qvB and is directed transverse to both \mathbf{v} and \mathbf{B} . It therefore moves in a circle with a certain radius ϱ , such that the magnetic force qvB is balanced by the centrifugal force mv^2/ϱ . From this it is found that the *gyro radius* ϱ is given by

$$\varrho = mv/qB \quad (1.3.5)$$

As the particle usually also has a velocity component parallel to the magnetic field, the resulting total motion describes a *spiral* on the surface of a circular cylinder with its axis parallel to the local magnetic field.

Thanks to a perturbation method, described in sections 3.7 and 4.8, this motion, too, can be described in a convenient way.

New Kinds of Waves

In the plasma state matter can carry many kinds of waves. The simplest of these are the *magnetohydrodynamic waves*, nowadays called *Alfvén waves* after their discoverer (Alfvén 1942).

These waves, which are possible because of the combination of deformability and electric conductivity, are the result of a coupling between the electromagnetic field and the motion of matter. The propagation speed that characterizes these waves, the *Alfvén velocity*, V_A , is in a plasma often more important than the velocity of sound. It is given by the expression

$$V_A = B / \sqrt{\mu_0 \rho_m} \quad (1.3.6)$$

where ρ_m is the mass density and B the magnetic field strength.

Another example of waves in a plasma is that of *whistler waves*, by means of which the first indications of plasma in outer space (beyond the ionosphere) were found (Chapter 4).

A theoretical analysis of waves in magnetized plasma, with whistler waves as a special case, is given in section 3.9.

“Frozen-in Magnetic Field Lines”

Another manifestation of the coupling between material motion and magnetic fields is the law of “*frozen-in magnetic field lines*”. This is valid if the electric conductivity is sufficiently high. What it implies is that the deformation of the medium and the change of the magnetic field follow each other in such a way that, if two elements of the medium are connected by a common field line at some instant, they are so connected at any other time, no matter how much the medium and the field have been deformed. This is illustrated in Fig. 1.3.2. (A derivation of the law and its conditions for validity are given in section 6.3.)

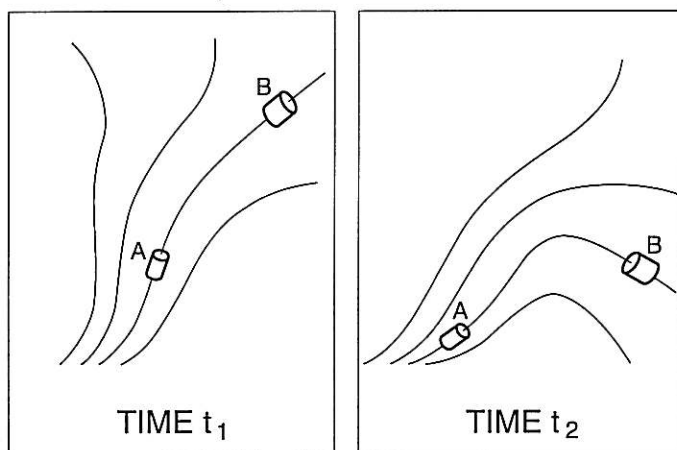


Figure 1.3.2. “Freezing-in” of magnetic field lines implies that elements of matter that are at some instant, t_1 , connected by a common magnetic field line, are so connected at any other instant of time, t_2 .

The law of frozen-in magnetic field lines leads to a drastically simplified description of many plasma physical phenomena. In space plasma the condition of validity is usually satisfied with a wide margin. But it has also been found that the condition for “freezing-in” can be violated by local electric fields along the magnetic field lines. Some of today’s most interesting scientific problems in the field of space plasma physics are related to such electric fields. They occur for example above the aurora (section 4.5) and are an important prerequisite for acceleration of the auroral primary particles (*i.e.* the particles, mostly electrons, which by impinging on the upper atmosphere cause the aurora).

Inhomogeneity

From laboratory experiments it is a well-known fact, that plasma carrying electric current has a pronounced tendency to develop *inhomogeneities*, especially *filamentary structures*. A consequence of this is that mathematical models that, for reasons of simplicity, assume a homogeneous plasma, easily become misleading.

There are many examples of cosmical plasmas that have a filamentary structure (comet tails, prominences, the solar corona, gas nebulae etc.) and the filamentary structure is often associated with electric currents.

A new kind of inhomogeneity, which was *not* foreseen from laboratory experiments but was only discovered from *in situ* measurements in space, is the remarkable tendency of the space plasma to form distinct regions, separated by surprisingly well-defined boundary surfaces, which often coincide with electric surface currents. Examples are the magnetopause (section 4.1) and the interplanetary current sheet (section 6.1). In other words, the space plasma has a pronounced tendency to form *cellular structures* (Alfvén 1981).

1.4. Regions of Space

The presentation in this book will follow a “geographic” sequence, beginning at the Earth. Thus the first topic will be the *atmosphere* (Chapter 2), although only its uppermost part can be considered to be part of outer space. Knowledge of the properties of the atmosphere is important as a background for study of the technical means of space research (balloons, sounding rockets and satellites).

The uppermost layers of the atmosphere are ionized, mainly by the short-wavelength part of the solar radiation (ultraviolet light, and soft X-rays), to some extent also by particle precipitation, including cosmic radiation. This ionized part of the atmosphere constitutes the *ionosphere* (Chapter 3). The ionosphere, which begins at less than 100 km above the surface of the Earth, constitutes the closest region of the cosmic plasma, which from there extends throughout the universe.

The region of space that is penetrated by the Earth’s magnetic field forms a rather well-defined volume, the *magnetosphere* (Chapter 4). Its lower boundary is the ionosphere, which by friction is tied to, and rotates with, the Earth’s atmosphere. The outer reaches of the magnetosphere are under the influence of a dynamic coupling with the interplanetary plasma.

The cause of the limited reach of the Earth's magnetic field is the *solar wind* (Chapter 6), which streams through the whole solar system with a speed of 300 – 900 km/s. As the solar wind also contains a “frozen-in” magnetic field, it acts as a so-called magnetohydrodynamic (MHD) generator, which drives electric currents of several mega-ampère through the magnetosphere to and from the Earth's atmosphere, causing auroras and geomagnetic storms. In front of the magnetosphere, which constitutes an obstacle to the (supersonic) solar wind, there is formed a *bow shock* analogous to the shock wave in front of a supersonic aircraft. This one, however, is a so-called “*collisionless*” *shock wave* – a never before observed physical phenomenon.

The source of the solar wind is the sun's outermost atmosphere, the corona. The *sun* (Chapter 5) is entirely a plasma. At its center the temperature is sufficiently high (about 15 million K) to sustain thermonuclear reactions, which release a power of about $4 \cdot 10^{26}$ W. By convection and radiative transport the energy reaches the surface, where most of it is emitted as radiation into space, and the temperature sinks to about 6 000 K. As a consequence of not yet fully understood processes, the temperature rises again outwards and in the corona it reaches about $2 \cdot 10^6$ K, which is enough to cause the omnidirectional outflow that becomes the solar wind.

The solar wind penetrates the entire *interplanetary space* (Chapter 6) and reaches far beyond it. How far is not known with certainty. The region of space engulfed by the solar wind is called the *heliosphere* (the sun's “magnetosphere”) and its outer border the *heliopause*.

Just as the Earth's magnetosphere is an obstacle in the flow of the solar wind, the same is true for the other planets in the solar system, and for satellites (moons) outside the magnetospheres of their host planets. Therefore, magnetized planets are surrounded by magnetospheres, which can, however, be very different from that of the Earth. At planets or moons without a magnetic field of their own, the solar wind interacts directly with the unprotected atmosphere (if there is one) or the solid surface (as in the case of the Earth's moon). *Magnetospheres and ionospheres* other than those of the Earth are also treated in Chapter 4.

Like the sun, all *stars* are in the plasma state, and if they are hot enough, they surround themselves with stellar winds. Even beyond the reach of the solar wind (and stellar winds) space is filled with an *interstellar plasma*, which, penetrated by a weak interstellar magnetic field, fills our entire galaxy, the Milky Way. The same is true for other galaxies. *Stellar and interstellar plasma* is treated in Chapter 7.

Even outside the galaxies there is a thin plasma. This *intergalactic plasma* (Chapter 8), is also magnetized and in some cases contains super-relativistic electrons.

The *cosmic radiation* can be considered as an extremely high energy subpopulation of the cosmic plasma. It is treated in Chapter 9.

A useful tool for clarifying the physical properties of the space plasma is to not only passively observe phenomena that take place spontaneously in space, but to also perform *active experiments*, that is to introduce controlled perturbations and observe the response of the system. Thus, one can utilize *space as a laboratory*. Some examples of this are given in Chapter 10.

2. The Atmosphere

The pictures in this chapter are, unless otherwise stated, from the *Handbook of Geophysics and the Space Environment* (1985).

2.1. Origin

The main ingredient of the atmosphere, nitrogen (78 % in the troposphere) originates from volcanic activity and denitrification of decaying organic substances. It is removed from the atmosphere by lightning discharges, certain photochemical processes and by the life processes of microbes. The oxygen (20.9 %) is produced mainly through the photosynthesis of plants and removed by oxidation processes, such as the decomposition of minerals, combustion and the respiration of organisms. Most of the atmospheric components undergo a circulation. The residence time in the atmosphere can be very different, for example a couple of weeks for water vapour and 100 000 years for argon (0.9 %). Helium and hydrogen, on the other hand, are lost into space as a non-negligible fraction of their atoms or molecules reach escape velocity (11.3 km/s).

2.2. The Standard Atmosphere

For use in aeronautical and space activities WMO (World Meteorological Organization) has defined a “*Standard Atmosphere*”. It is described as “A hypothetical vertical distribution of atmospheric temperature, pressure and density, which by international agreement and for historical reasons, is roughly representative of year-round midlatitude conditions. Typical usages are as a basis for pressure altimeter calibrations, aircraft performance calculations, aircraft and rocket design, ballistic tables and meteorological diagrams. The air is assumed to obey the perfect gas law, and the hydrostatic equation, which, taken together, relate temperature, pressure and density with geopotential.”

The standard atmosphere in use at present is *USSA76* (U S Standard Atmosphere 1976). For rocket and satellite applications COESA (Committee on Extension to the Standard Atmosphere) has made the following addition:

“The atmosphere shall also be considered to rotate with the Earth, and be an average over the diurnal cycle, semiannual variation, and the range of conditions from active to quiet geomagnetic, and active to quiet sunspot conditions. Above the turbopause (about 110 km) generalized forms of the hydrostatic equations apply.”

It is important to remember that the *real* atmosphere varies with time and place and therefore always deviates more or less from the Standard Atmosphere.

The *density* and *pressure* of the Standard Atmosphere decrease with altitude as shown in Fig. 2.1. The *temperature* variation up to 100 km altitude is shown

in Fig. 2.2. There are also marked the atmospheric layers, which are described in more detail below. As can be seen in the figure, the temperature distribution at high altitudes is different in different years. This is due to differences in the solar radiation, which varies with the 11-year *solar cycle* (Chapter 5).

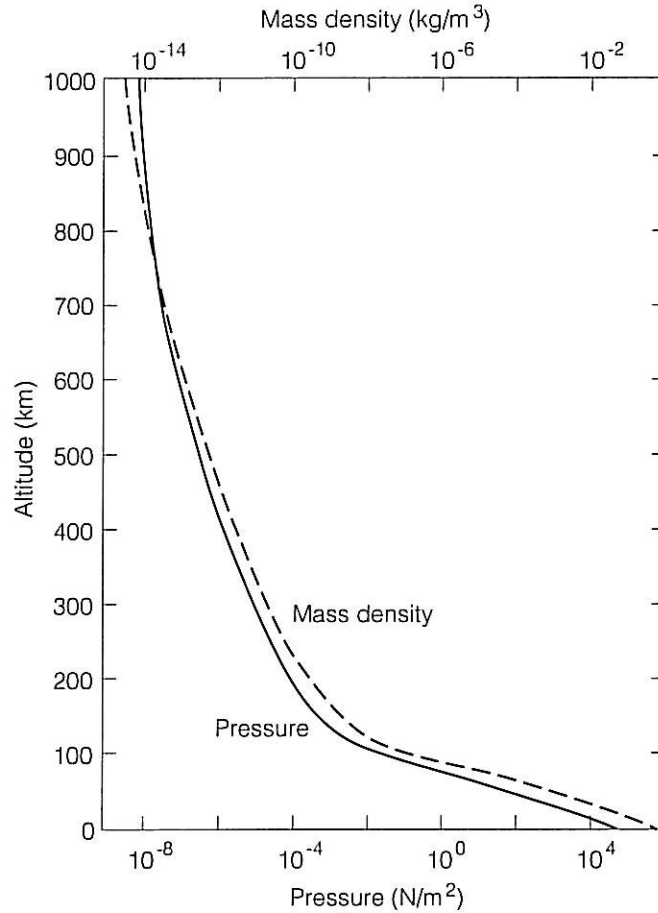


Figure 2.1. Altitude variation of pressure in the Standard Atmosphere.

The solar cycle variations are even more pronounced at higher altitudes. This is clearly evident in Fig. 2.3, which shows the temperature distribution in the interval up to 300 km for years with, respectively, low and high solar activity. The years shown in the figure are, however, no extreme cases. The temperature above 200 km can vary from 600 K at solar minimum to 2 000 K at solar maximum.

What is shown in Fig. 2.3 is the “kinetic” temperature. At high altitudes (above about 70 km) the atmosphere is not in local thermodynamic equilibrium, (LTE). It is appropriate to use the concept of *kinetic temperature*, which is defined in terms of the mean energy of the air molecules. A kinetic temperature T means that the mean energy of the molecules is the same as it would have been, if the air had been in thermodynamic equilibrium at the temperature T , that is $3kT/2$. Thus the root mean square velocity is

$$v_{rms} = \sqrt{\frac{3kT}{m}} \quad (2.2.1a)$$

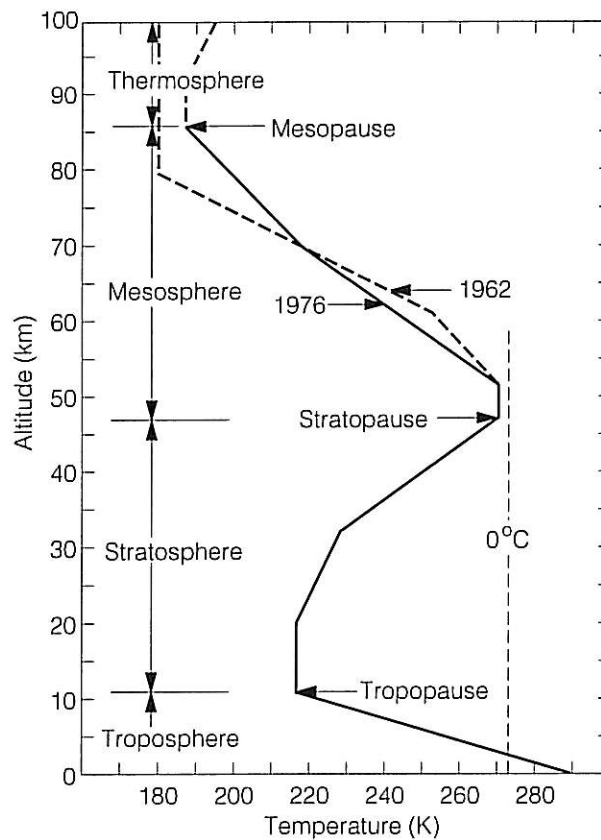


Figure 2.2. Altitude variation of temperature in the Standard Atmosphere in the altitude interval 0–100 km.

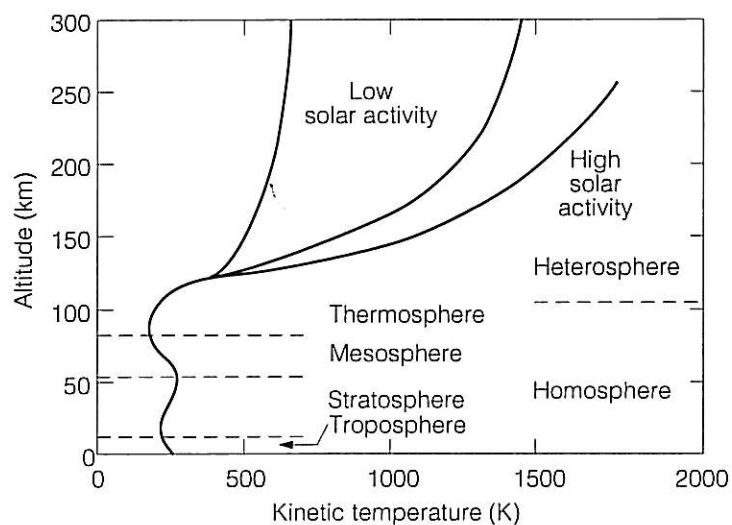


Figure 2.3. Altitude variation of temperature in the Standard Atmosphere, altitude interval 0–300 km.

where m is the mean molecular mass and k is Boltzmann's constant, $k = 1.38 \cdot 10^{-23} \text{ J/K}$. In the context of collision and transport processes the more relevant quantity is the *linear* average velocity

$$\bar{v}_T = \sqrt{\frac{8kT}{\pi m}} \quad (2.2.1b)$$

These velocities are of the same order of magnitude as the *sound speed*, which is given by

$$V_s = \sqrt{\frac{\gamma kT}{m}} \quad (2.2.2)$$

where $\gamma = 1.4$ is the ratio between the air's specific heat at constant pressure and at constant volume.

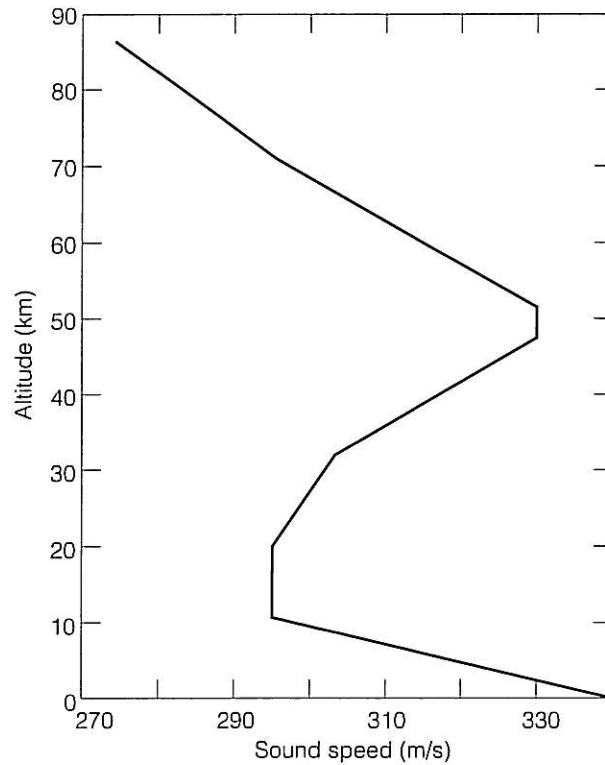


Figure 2.4. Variation of sound speed in the Standard Atmosphere, interval 0–90 km.

The altitude variation of the sound speed up to about 100 km, Fig. 2.4, is therefore very similar to that of the temperature, because the mean molecular mass is nearly constant. On the other hand, the latter decreases rapidly at altitudes above 110 km, as Fig. 2.5 shows. This, of course, reflects the composition of the atmosphere, which is shown in Fig. 2.6 for the altitude range 0 – 1000 km. Below about 110 km the composition of the atmosphere is rather homogeneous as a result of turbulent motion mixing the ingredients. Therefore the region below about 110 km is called the *homosphere*, the region above is called the *heterosphere* and the boundary between them the *turbopause*. In the heterosphere molecular oxygen and nitrogen decrease, so that atomic oxygen dominates in the 200 km to 600 km altitude interval. Above that atomic hydrogen and helium take over.

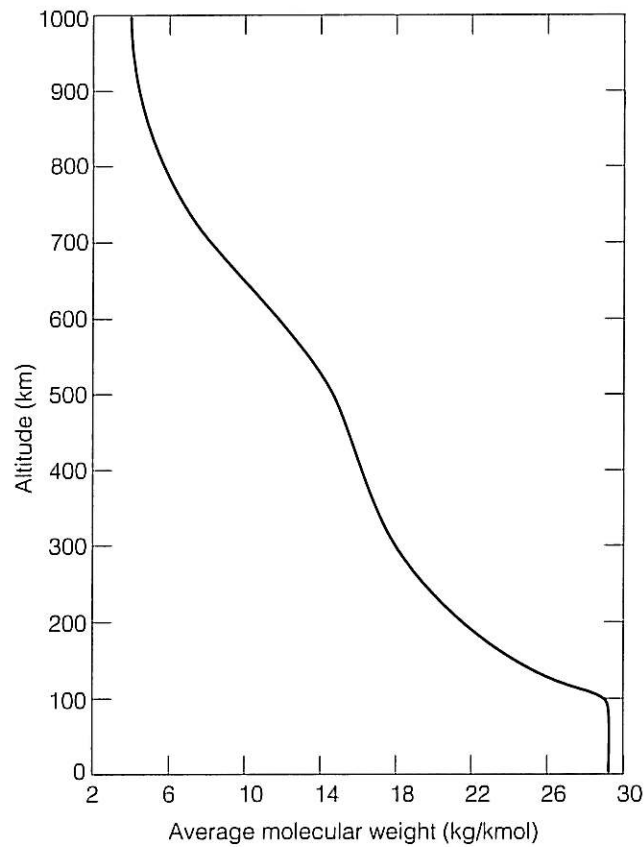


Figure 2.5. Mean molecular mass in the Standard Atmosphere as a function of altitude.

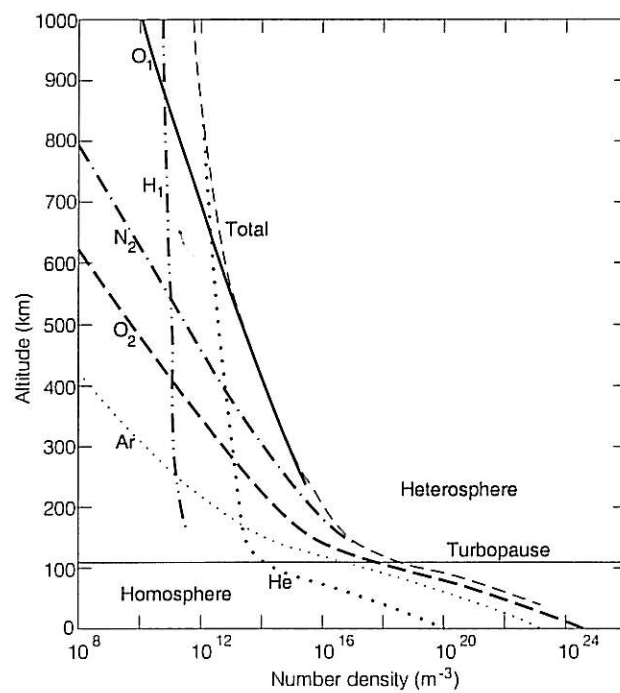


Figure 2.6. Altitude distribution of the main components of the Standard Atmosphere.

2.3. Scale Height

In hydrostatic equilibrium the gravitational force, which for a volume element with density ρ_m , is equal to $g\rho_m$ and directed downward, is balanced by the pressure force $-dp/dz$ (where g is the acceleration of gravity, 9.8 m/s^2 and z is the altitude coordinate, positive upward). Thus

$$-\frac{dp}{dz} = g\rho_m \quad (2.3.1)$$

The pressure is related to the absolute temperature T by the general gas law, $p = RT\rho_m/M$, where R is the general gas constant ($8.314 \text{ J/mol}\cdot\text{K}$) and M the molar mass (that is, the mass of one mol of the substance, which is the same as the molecular mass, m , times Avogadro's constant, $A = 6.02217 \cdot 10^{23}$ particles per mol). If the temperature were constant, $-dp/dz = g\rho_m = -(RT/M)d\rho_m/dz$, and thus $d\rho_m/dz = -(gM/RT)\rho_m$, from which it follows by integration that

$$\rho_m = \text{const. } e^{-z/(RT/gM)} = \text{const. } e^{-z/H} \quad (2.3.2)$$

where $H = RT/gM$ has the dimension of length and represents the altitude increment that is required for the density to decrease by a factor e . The parameter H is called *scale height*. As $R = Ak$, and $M = Am$, where k is Boltzmann's constant ($k = 1.38 \cdot 10^{-23} \text{ J/K}$), and m is the molecular mass, Avogadro's number cancels out, and the following alternative expression is found for the scale height

$$H = RT/gM = kT/gm \quad (2.3.3)$$

In the homosphere the mean molar mass is 0.029 kg/mol . At a temperature of 293 K (20° C) the scale height becomes about 10 km . As the temperature varies strongly with altitude, the real large-scale distribution is far from exponential, and the concept of scale height can only be applied *locally*.

Above the turbopause the components of the atmosphere assume individual height distributions determined by diffusive equilibrium. Each one has its own scale height, $H_x = RT/gM_x$, where x designates the molecular species (oxygen, nitrogen, etc.)

Knowledge about the higher atmospheric layers has of course increased greatly as a result of space research. An early method of investigating the upper atmosphere was the so-called sound grenade method. From a high altitude rocket a number of explosive packages were ejected and detonated at different altitudes. Differences in the arrival time at the ground of the sound from the different explosions gave information about the sound speed in the height interval concerned.

The density of air at high altitudes is of course important for the life span of satellites. An early example is the Echo-satellite, an electrically conducting balloon constituting a passive radio wave reflector. Its life span turned out to be considerably shorter than expected, the reason being that the density of helium at high altitudes had been underestimated. We now know that, typically, the altitude region $1000 - 3000 \text{ km}$ is dominated by helium and above that by hydrogen (Fig. 2.7), but great

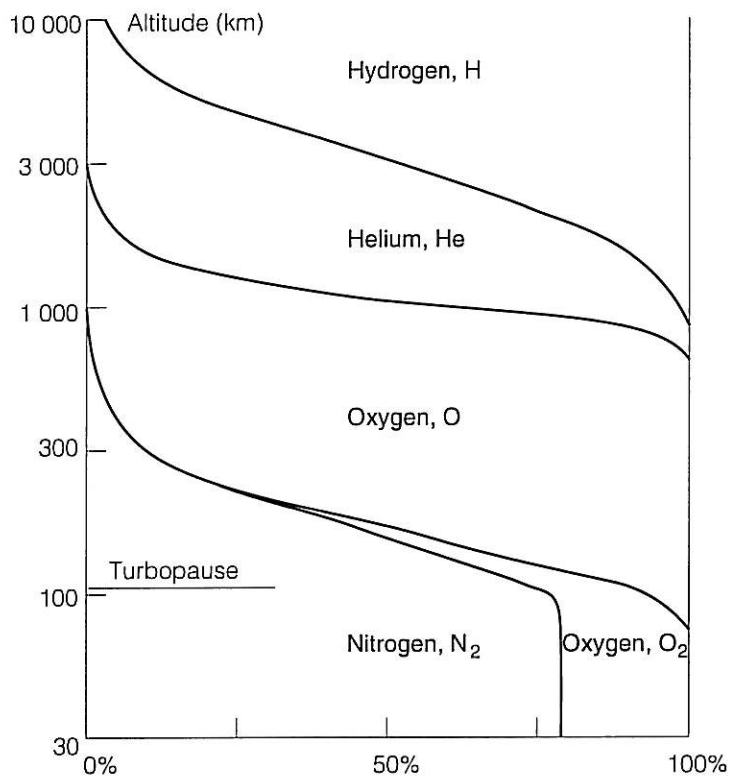


Figure 2.7. Relative distribution of the main components of the atmosphere in the altitude range 30 – 10 000 km (Bolin 1969).

variations occur in the altitude distributions. The hydrogen-dominated uppermost part of the atmosphere is sometimes called the *geocorona*.

2.4. Structure of the Atmosphere

The atmosphere is usually divided into a number of different layers characterized by different physical conditions. Above has already been mentioned a division based on *transport processes*, namely turbulent mixing in the homosphere and diffusive equilibrium in the heterosphere.

The most important division is based on the *thermal* conditions. The corresponding atmospheric layers, which are marked in Fig. 2.2, are the following.

The region where weather takes place extends from ground level to about 10 km altitude and is called the *troposphere*. As Fig. 2.2 shows, the temperature decreases about linearly with altitude, about 7°C per kilometer to about –60° C at the upper boundary of the troposphere, the *tropopause*. Large deviations from this mean variation occur, for example depending on the content of water vapour at different altitudes.

In the next layer, the *stratosphere*, the temperature is first constant in a layer of about 10 km thickness and then rises until the upper boundary of the stratosphere, the *stratopause*, at an altitude of about 45 km. In the next higher layer, the *mesosphere*, the temperature falls again and at its upper boundary, the *mesopause*,

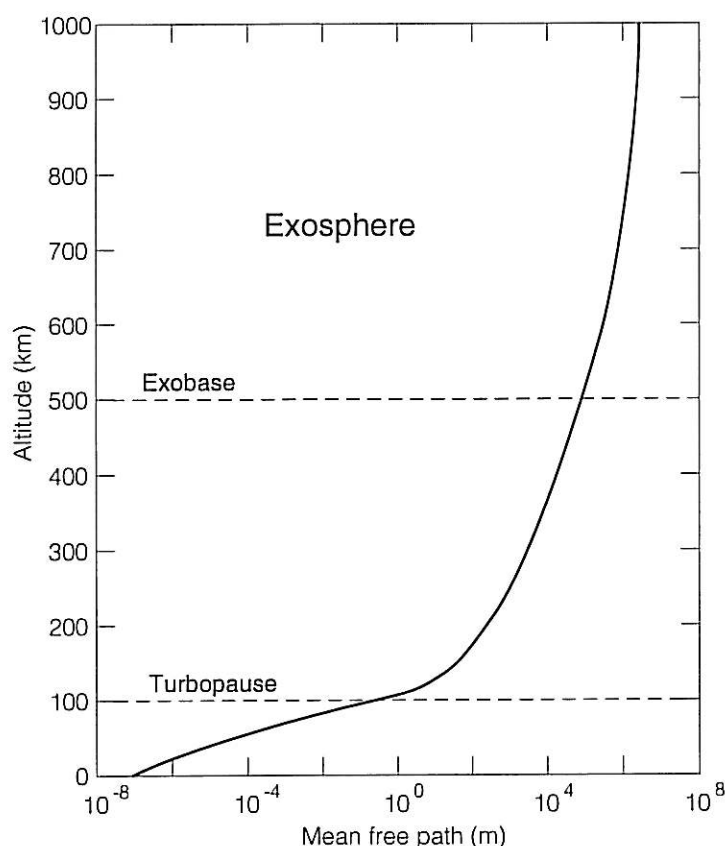


Figure 2.8. Altitude variation of the mean free path in the Standard Atmosphere.

it reaches its lowest value, about -90°C . At even greater altitudes, in the *thermosphere*, the temperature rises again and now to much higher values, as shown in Fig. 2.3. In this region no thermodynamic equilibrium prevails, and the temperatures quoted are the previously defined “kinetic” temperatures. As the density of the atmosphere decreases, the collisions between the air molecules become increasingly rare. The distance, the *mean free path*, that a molecule in average moves between two successive collisions, increases in the way shown in Fig. 2.8.

At sufficiently high altitudes the collisions become so few that the atoms and molecules move in unperturbed ballistic orbits. This outermost region of the atmosphere is called the *exosphere*. Its lower boundary, the *exobase*, is located at an altitude of about 500 – 1000 km.

For most of the particles the ballistic orbits are segments of ellipses with the center of the Earth at one of the foci, and the particle falls repeatedly below the exobase and undergoes a collision. But in a few cases these collisions give the particle a speed exceeding the *escape velocity*, that is the minimum speed required to leave the gravitational field of the Earth (11.3 km/s). The orbit then becomes a branch of a hyperbola, and the particle escapes into space.

As an example of how the real atmosphere can deviate from the Standard Atmosphere, Fig. 2.9 shows the result of a rocket experiment, “CAMP” for determination of the temperature around the mesopause.

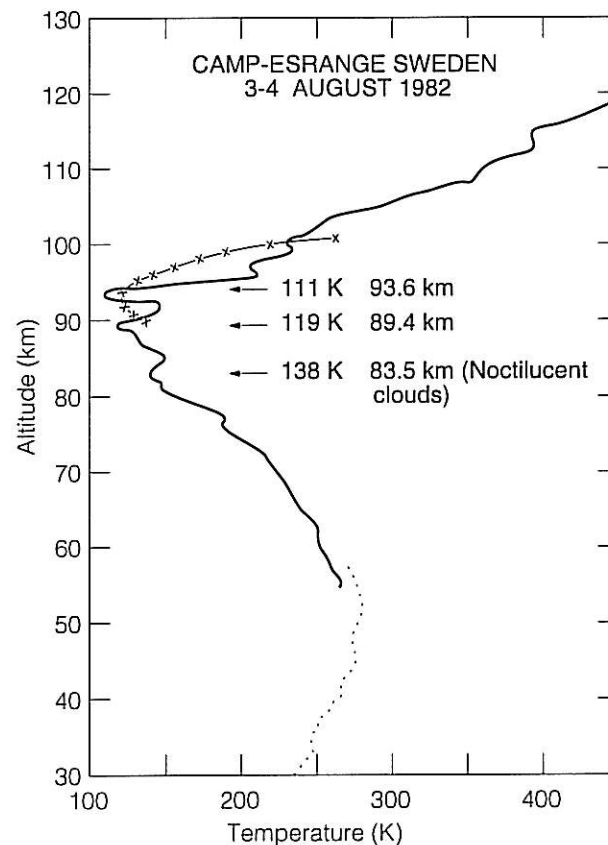


Figure 2.9. Temperature profile measured in the experiment "CAMP". Note the extremely low temperatures at the mesopause, which deviate greatly from the temperature values of the Standard Atmosphere (Philbrick *et al.* 1984).

2.5. Solar Radiation

The energy flux from the Sun that reaches the uppermost atmosphere of the Earth is called the Sun's (total) *irradiance*. Its average value, corrected for seasonal variations in the Sun–Earth distance, is called the *solar constant*. Its value has been determined to be $1373 \pm 20 \text{ W/m}^2$ (Frohlich, 1977). How the energy is distributed between different parts of the solar spectrum is shown in Fig. 2.10, which shows the irradiance as a function of wavelength. (For environmental tests of space vehicles this spectrum is simulated in test chambers by means of suitable artificial light sources.)

The *Solar Cycle*, which is described in Chapter 5, influences the Sun's irradiance mainly in the ultraviolet region (Fig. 2.11), where a relatively small part of the total energy is radiated. As the UV radiation is absorbed already in the upper atmosphere, it is mainly the upper atmospheric layers that are influenced by the solar cycle, as can be seen from Fig. 2.3.

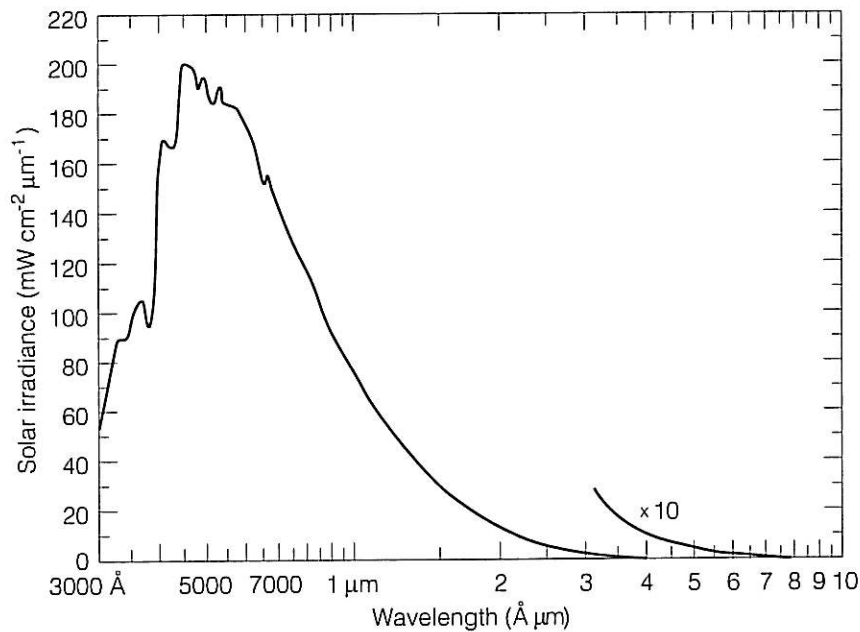


Figure 2.10. The solar irradiance at the upper boundary of the Earth's atmosphere as a function of the wavelength (Pierce and Allen, 1977).

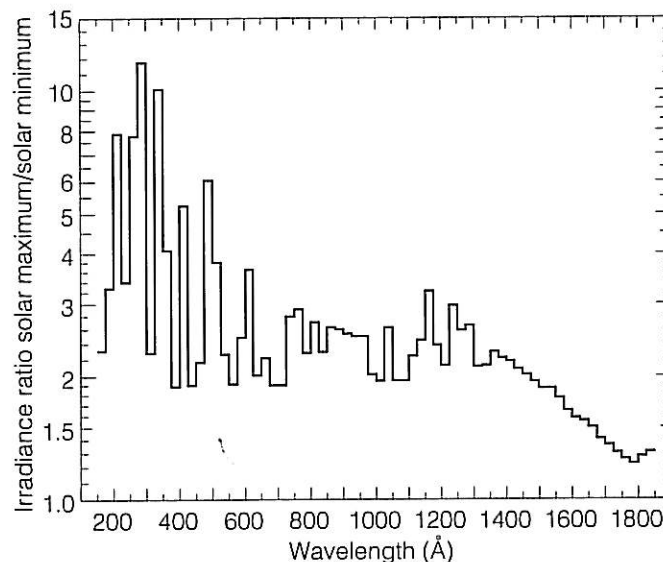


Figure 2.11. Ratio of irradiance at high and low solar activity as a function of wavelength in the ultraviolet part of the spectrum (Hinteregger, 1981).

2.6. Electric Conductivity

The electric conductivity of the atmosphere, which is very low at the surface of the Earth ($10^{-14} - 10^{-13}$ S/m) grows with altitude. In the altitude interval 10 – 90 km the variation is approximately exponential for six powers of ten with a scale height of 7 – 8 km (Hale *et al.* 1981). A consequence of this altitude distribution of atmospheric conductivity is that large-scale horizontal electric fields in the ionosphere penetrate with little attenuation to an altitude of about 30 km.

This means that balloon-borne electrometers can be used to measure the large-scale horizontal electric field in the ionosphere, which has been used in international research projects with balloon launches from Scandinavia and from Antarctica. An empirical model of atmospheric conductivity up to 40 km altitude is given in Fig. 2.12. Of course, the actual distribution of conductivity at a given time or place varies. But even large deviations from the exponential distribution cause only small errors in the determination of the ionospheric electric field, provided the balloons go high enough.

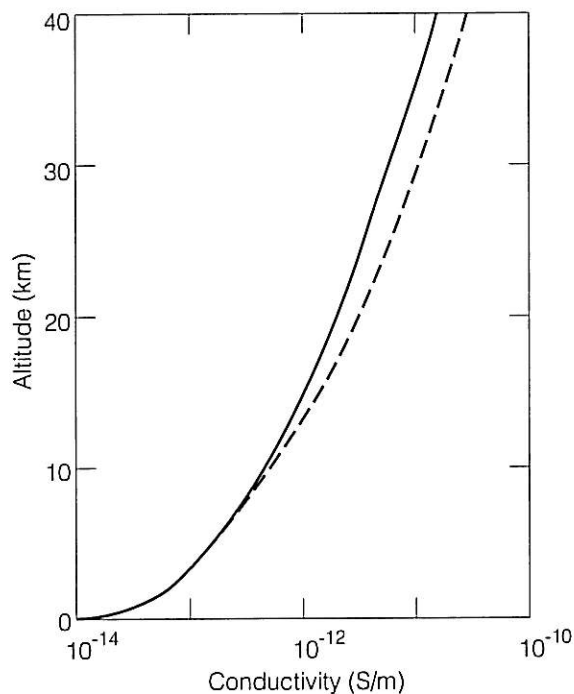


Figure 2.12. Electric conductivity of the atmosphere in the altitude interval 0–40 km, empirical model (Makino and Ogawa 1984). The solid and dashed curves apply to latitudes of 0 and 40, respectively.

3. The Ionosphere

3.1. History

The existence of an electrically conducting layer in the upper atmosphere was suggested in 1883 (by Stewart) in order to explain the daily variations of the Earth's magnetic field. In 1902 Kennelly and Heaviside used the same hypothesis to explain Marconi's successful experiments with transatlantic radio communication. After the validity of the hypothesis had been proved experimentally by Appleton and Barnett by means of reflection of radio waves, this conducting layer was for a long time called the *Kennelly-Heaviside-layer*.

The main reason why the upper atmosphere is electrically conducting is that the radiation of the sun at wavelengths less than $2\,400\text{ \AA}$ (ultraviolet and soft X-rays) ionizes air molecules. To some extent particle radiation from outer space also contributes. Particle radiation with energies of the order of a few keV is especially important as the direct cause of the aurora (Chapter 4.5). The so-called *cosmic radiation* (Chapter 9) with extremely high energy is responsible for ionization in the lowest altitude range, about 70 km.

The electron density of the ionosphere has a maximum at an altitude of about 250 km. The formation of a maximum can be understood by the following qualitative argument. At very high altitude the solar radiation has its full intensity, but the density of the atmosphere is very small. Thus there are few particles to ionize, and the electron density becomes small. With decreasing altitude there are more and more atmospheric particles for the radiation to ionize, and therefore the electron density increases. But it also means that the intensity of the radiation decreases as more and more of it is absorbed. When the decrease of intensity just balances the increase of atmospheric density, a maximum is reached, and from there the electron density decreases with decreasing altitude until all of the ionizing radiation is absorbed. A quantitative calculation, given in a later section (3.5), leads to a characteristic electron density profile called the *Chapman distribution*.

In reality the electron density distribution becomes more complicated because different molecules are ionized at different altitudes. As a result, the ionosphere can be divided into a number of layers, as discussed below (in 3.2).

Even before the Space Age the ionosphere was studied by means of ground based *ionosondes*, radio transmitters whose signals were reflected by the ionospheric layers. A radio wave with a certain frequency, f , is reflected at an altitude, where the electron density reaches a certain value (as described in section 3.4). The higher the transmitted frequency, the higher density, and the higher altitude the wave can reach before being reflected. By measuring the time delay of the reflected signals as a function of frequency it is thus possible to calculate the electron density as a function of altitude.

It is customary to represent the reflected signal in a so-called *ionogram*, where the frequency, which is a measure of the electron density at the reflection level, is plotted on the horizontal axis and the delay, which is a measure of the altitude

of the reflection level, on the vertical axis. The strength of the reflected signal is represented by the degree of blackening. An example of an ionogram is given in Fig. 3.1.1.

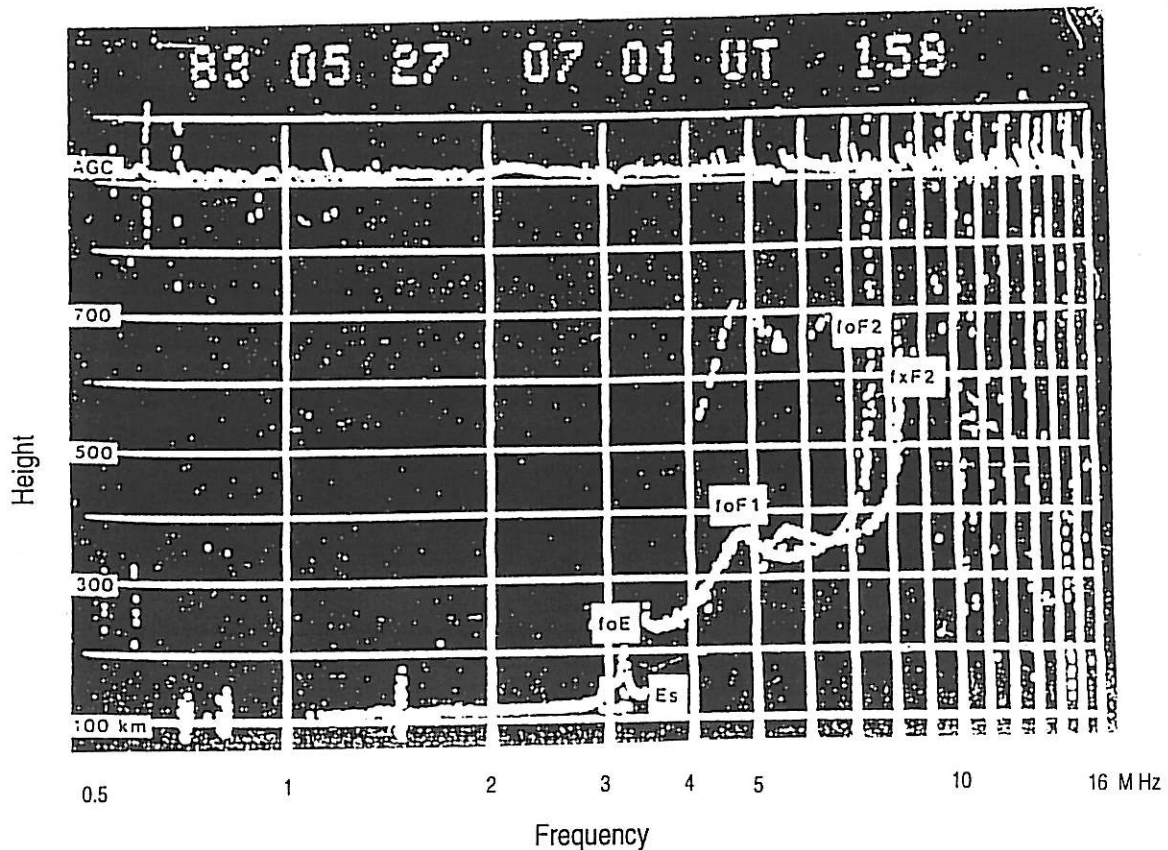


Figure 3.1.1. Example of an ionogram.

This method can only give information about the electron density below the density maximum, because signals with frequencies corresponding to the topside electron density never get there because they are reflected below. It is therefore only by means of space technology that it has become possible to study the electron density distribution in the topside ionosphere. This was first done by sounding the ionosphere from above by means of satellite-borne ionosondes using the same principle as the ground-based ones.

Nowadays it is, however, possible to study the whole ionosphere even from the ground using a technique called *incoherent scattering* or *Thomson scattering*. This technique makes use of the fact that each individual charged particle scatters radio waves. As the effective scattering cross section is small (for an electron it is $8\pi r_e^2/3$, where r_e is the classical electron radius $r_e = e^2/4\pi\epsilon_0 m_e c^2 = 2.82 \cdot 10^{-15}$ m), the reflected signal is very weak, so that large transmitter and receiver antennas are needed. The first facility for probing the ionosphere by means of incoherent scattering was built at Jicamarca in Peru in the early 1960s (with a transmitted power of 6 MW and 9 216 receiver dipoles distributed over an area of 84 000 m²). Nowadays there are several incoherent scatter facilities around the world, including one in Northern Scandinavia, namely *EISCAT* (*European Incoherent Scatter Scientific Association*), with ground stations in Kiruna, Tromsø and Sodankylä.

This technique is very powerful, and from the spectrum of the reflected wave it is possible to determine not only electron density but also ion temperature, ion species, velocity of the plasma, ion collision frequency and strength of the magnetic field. A description of the technique has been given by Folkestad (1979,1981).

3.2. The Ionospheric Layers

The actual structure of the ionosphere is the result of an interplay between the spectral distribution of the incident radiation and the occurrence of different molecular species in the upper atmosphere. The positive charges that neutralize the electrons are therefore carried by different kinds of positive ions at different altitudes. It is customary to distinguish four different ionospheric layers, D, E, F1 and F2. Sometimes the lowest part of the D layer, below 70 km altitude, which is mainly caused by cosmic radiation, is counted as a separate layer, the C layer. Table 3.2.1 summarizes the typical altitudes, electron densities, ion contents and dominating causes of ionization. Although positive ions dominate, the lower part of the ionosphere also contains negative ions (such as NO_3^- , NO_2^- , O_2^- and CO_4^-). For example, Krankowsky *et al.* measured NO_3^- and CO_3^- densities of the order of 10 cm^{-3} .

Layer	D	E	F ₁	F ₂
Altitude (km)	60-85	85-140	140-200	200 - ca 1500
Nighttime electron density (cm^{-3})	$<10^2$	$2 \cdot 10^3$	—	$2 - 5 \cdot 10^5$
Daytime electron density (cm^{-3})	10^3	$1 - 2 \cdot 10^5$	$2 - 5 \cdot 10^5$	$0.5 - 2 \cdot 10^6$
Ion species	$\text{NO}^+ \text{O}_2^+$	$\text{NO}^+ \text{O}_2^+$	$\text{NO}^+ \text{O}_2^+ \text{O}^+$	$\text{O}^+ \text{He}^+ \text{H}^+$
Cause of ionization	Lyman α (1215 Å)	Lyman β (1025 Å) X-rays	UV	UV

Table 3.2.1. The ionospheric layers.

Typical electron density profiles for day and night, respectively, are shown in Fig. 3.2.1. As mentioned in section 2.5, the sun's shortwave radiation varies considerably with the so-called solar cycle (Chapter 5). How the ionosphere is influenced by this is also shown in Fig. 3.2.1.

As the most important cause of ionization, the solar radiation, only is present in daytime, the electron density distribution also becomes different in daytime and at night, as can be seen in the differences between Figures 3.2.1 a and b.

As seen in Fig. 3.2.1b, most of the D, E and F1 layers disappears at night. As the D layer, where the collision frequency, and therefore the absorption of radio waves, is relatively high, disappears at night, at the same time as the F2 layer largely remains, long distance shortwave communications over large distances (by repeated reflections against the ionosphere and the ground) are favoured at night. (The D

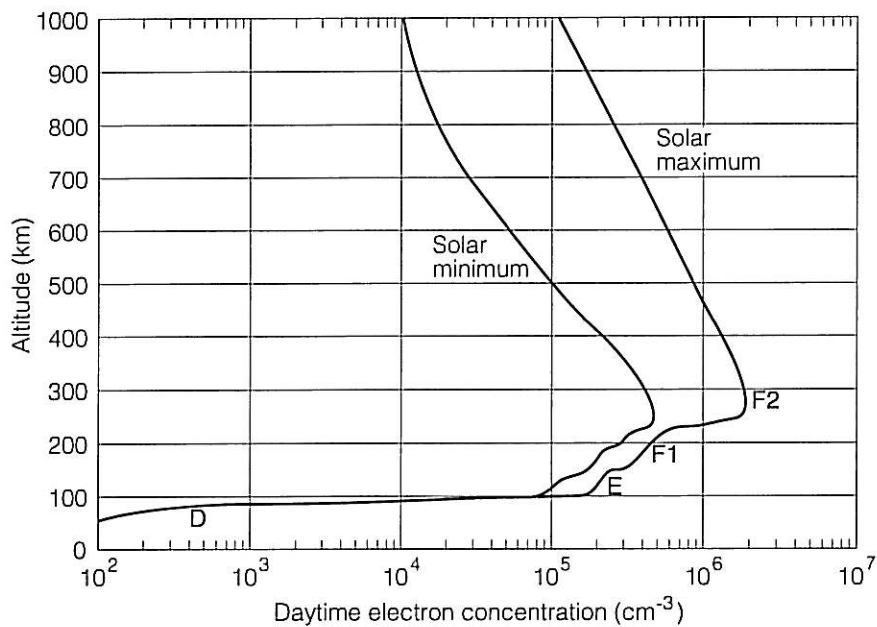


Figure 3.2.1a. Altitude distribution of electron density in the ionosphere in the daytime (Hanson 1965).

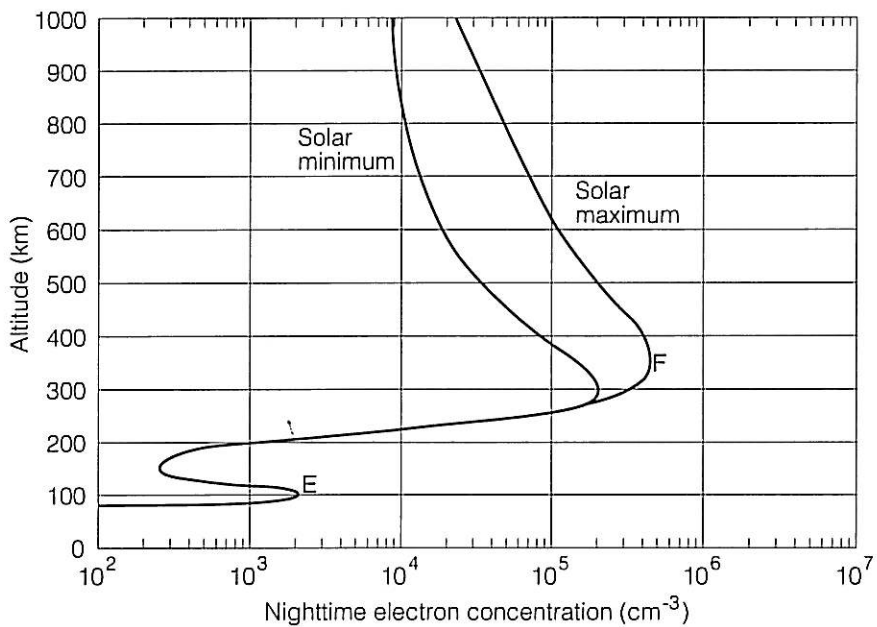


Figure 3.2.1b. Altitude distribution of electron density in the ionosphere at night (Hanson 1965).

layer can, however, also act as a reflector, namely for obliquely incident long radio waves, which, because of their low frequency, do not penetrate deep into the layer.)

As the density of the background atmosphere (and hence the collision frequency of electrons against neutrals) is very different in the different ionospheric layers, these have very different properties in terms of, for example, electrical conductivity and

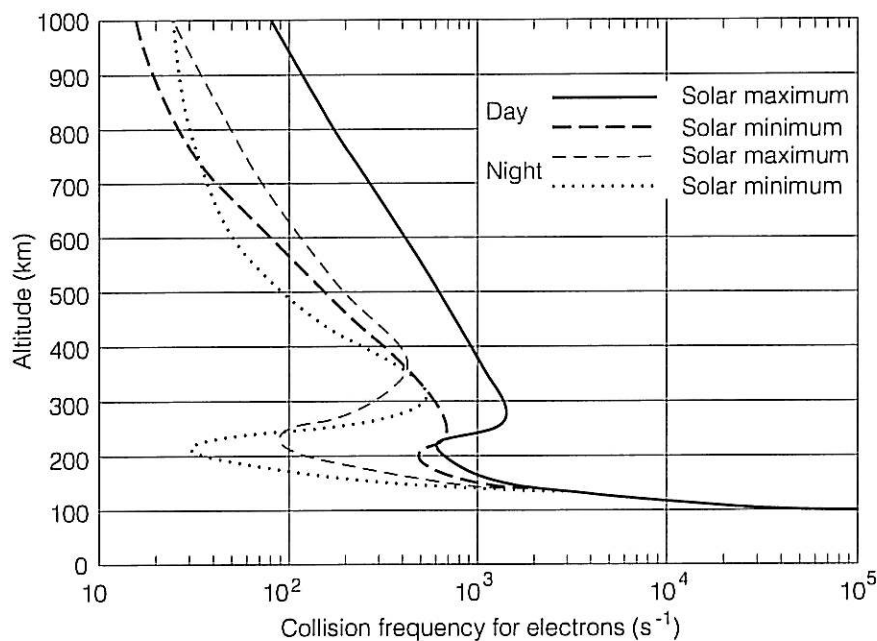


Figure 3.2.2. Altitude variation of the collision frequency of electrons in the ionosphere (Hanson 1965).

propagation of radio waves. Fig. 3.2.2 shows the altitude variation of collision frequency for daytime and nighttime at high and low solar activity.

Especially in the upper ionospheric layers the electrical conductivity is very different in different directions relative to the magnetic field. In other words, the ionospheric plasma at high altitudes has a very *anisotropic* conductivity. An electric field in one direction usually drives an electric current in a different direction. One must therefore distinguish between three different electrical conductivities. They are defined as follows (cf. Fig. 3.2.3)

1. *Parallel conductivity*, σ_{\parallel} , i.e. the conductivity in the direction parallel to the magnetic field.
2. The *Pedersen conductivity*, σ_P , the conductivity parallel to the component of the electric field that is perpendicular to the magnetic field.
3. The *Hall conductivity*, σ_H , the conductivity perpendicular to both the electric and magnetic fields (counted positive in the direction opposite to the vector product $\mathbf{E} \times \mathbf{B}$).

A theoretical analysis of the electrical conductivity is given in section 3.8. There it is also shown that the three conductivities can be regarded as components of a *conductivity tensor*.

The three conductivities depend on both electron and ion collision frequencies, and their values as functions of altitude are shown in Fig. 3.2.4. Note that in the upper layers, the parallel conductivity is several powers of ten greater than the others, which are of mutually comparable magnitude.

The ionosphere merges without a sharp boundary into the magnetosphere. It is therefore a matter of definition where to consider the ionosphere to end. A common definition, used in Table 3.2.1, is that it ends at a level of 1 000 – 2 000 km

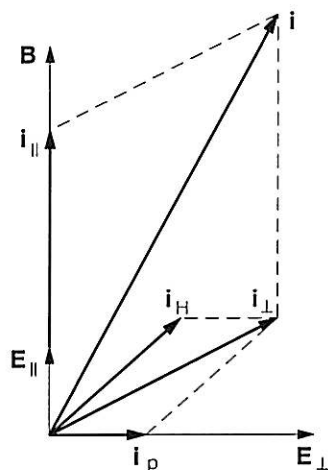


Figure 3.2.3. Definition of Pedersen, Hall and parallel current.

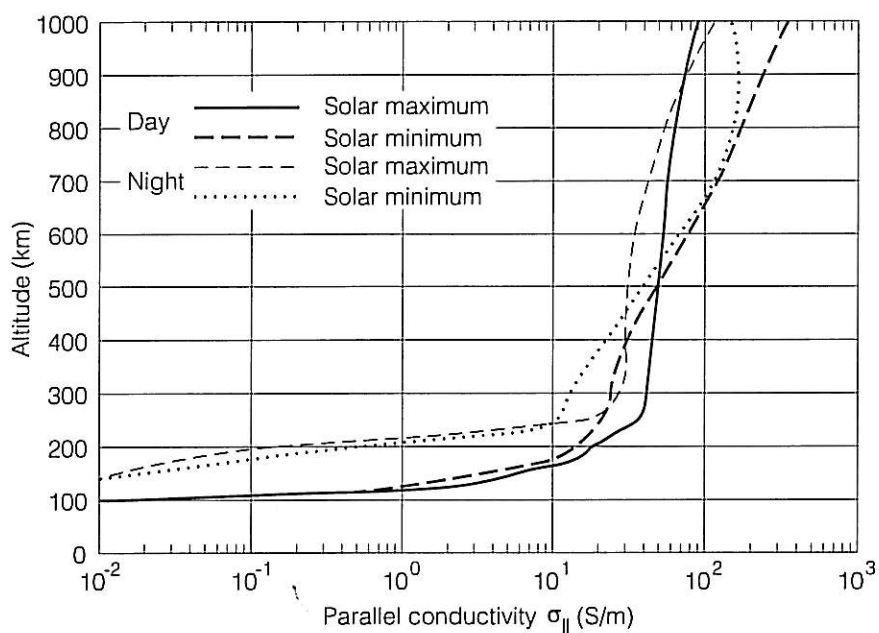


Figure 3.2.4a. Altitude variation of the parallel conductivity in the ionosphere (Hanson 1965).

altitude, above which the hydrogen ion is the dominating ion species during average conditions (from which great deviations occur).

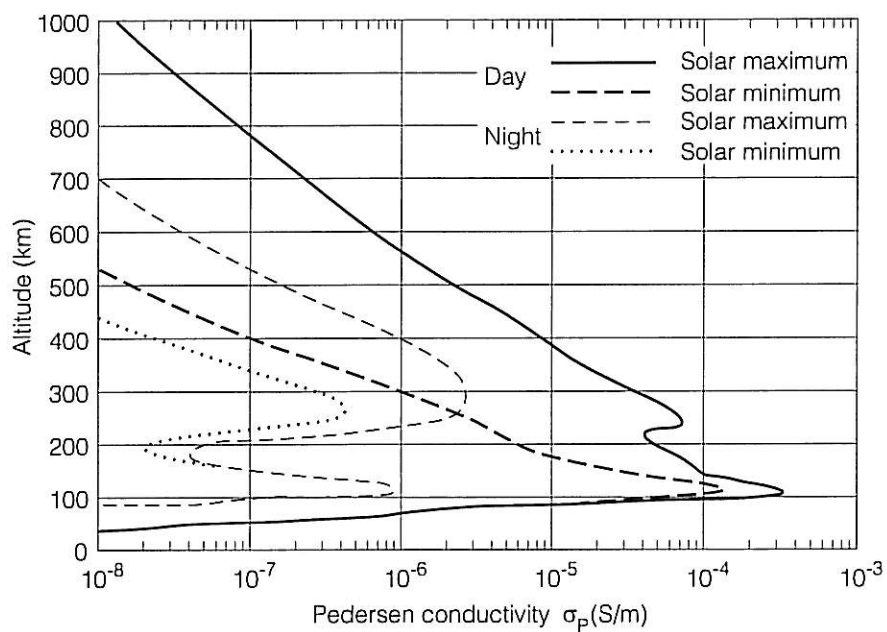


Figure 3.2.4b. Altitude variation of the Pedersen conductivity in the ionosphere (Hanson 1965).

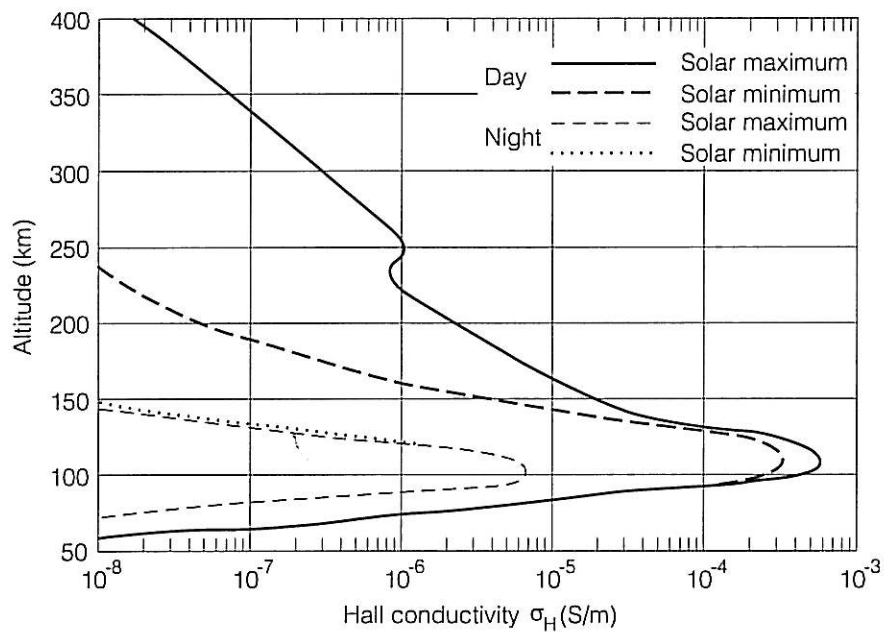


Figure 3.2.4c. Altitude variation of the Hall conductivity in the ionosphere (Hanson 1965).

3.3. Ionospheric Disturbances

Various kinds of *solar activity*, especially so-called *solar flares*, which are described in Chapter 5, lead to enhanced radiation, which can cause strong *ionospheric disturbances*. These disturbances occur with different time delays relative to the solar flare depending on which product of the flare is the cause. They can be divided into three principal categories described below.

Principal Category I: Sudden Ionospheric Disturbances, SID

This type of disturbance is caused by an intensification of the short wave electromagnetic radiation of the sun and therefore starts at the same time as the visible light reaches the Earth. The duration of the disturbance can vary from a few minutes to about an hour. The four most important disturbances of this category are:

- 1) Sudden Short-Wave Fadeout, SSWF
Radio reception in the frequency range 5–20 MHz (15–60 m wavelength) becomes difficult or impossible. The reason is enhanced ionization in the D layer, which leads to abnormally large absorption, because the collision frequency is high in this layer (cf. the section on radio wave propagation, 3.4).
- 2) Sudden Cosmic Noise Absorption, SCNA
The radio noise from outer space, which is routinely recorded in the wavelength range 6–17 m, is reduced due to the increased absorption in the D layer.
- 3) Sudden Phase Anomaly, SPA
Radio signals in the long wave band with oblique incidence can, for reasons mentioned in section 3.2, be reflected from the D layer. A sudden increase of ionization in the lowest ionospheric layers causes an equally sudden change of the reflection altitude and hence of the path length and the phase of the carrier wave. This is registered as part of the routine recording of the state of the ionosphere.
- 4) Sudden Enhancement of Atmospherics, SEA
Atmospherics in the long wavelength band caused by lightning discharges make themselves felt at greater distances because of the increased reflexion capability of the D layer for obliquely incident long wavelength waves. This is manifested in a higher level of atmospherics. Routine registration is usually done at wavelengths of about 10 km.

Principal Category II: Polar Cap Absorption, PCA

This kind of disturbances is caused by particle radiation from the solar flare. The particles are mainly protons and have energies in the range from a few MeV to

hundreds of MeV. They reach the Earth about 24 hours after the solar flare. Unlike the electromagnetic radiation, the particles are influenced by the Earth's magnetic field, and a consequence of this is that they can reach the ionosphere only in certain regions around the magnetic poles.

Where they do arrive, these particles cause ionization at very low altitudes, even down to 50 km, and cause a heavy absorption of radio waves, except for extremely large wavelengths, 10–30 km, whose propagation is instead enhanced. In spite of the fact that the solar flare itself has a short duration, (of the order of 15 minutes, cf. Chapter 5) and in spite of the fact that an ion pair created at these low altitudes recombines very rapidly, the disturbed state of the ionosphere can prevail for several days, in exceptional cases even a week. The reason is believed to be that the cloud of energetic particles is trapped in the interplanetary magnetic field and leaks out only slowly.

Principal Category III: Ionospheric Disturbances Associated with Geomagnetic Storms

In addition to electromagnetic radiation and high energy particles the solar flare also produces a stream of magnetized plasma, which, with a speed of nearly 1 000 km per second flows through interplanetary space – a kind of intensified solar wind (cf. Chapter 6). When this plasma reaches the vicinity of the Earth, a complicated interaction with the Earth's magnetic field takes place, which, among other things, leads to magnetic storms and enhanced auroras as described in section 4.5. The auroras cause a strongly enhanced local ionization in those regions of the ionosphere that are hit by the auroral particle fluxes.

3.4. Radio Wave Propagation in the Ionosphere

In this section is given a brief orientation. A detailed analysis is given in section 3.9.

Terminology

The terminology for radio wave frequencies is summarized in Table 3.4.1.

Radio wave bands

Ferequency band	Wavelength	Notation
20 Hz - 3 kHz	> 100 km	ELF (Extremely Low Frequency)
3 kHz - 30 kHz	100 km - 10 km	VLF (Very Low Frequency)
30 kHz - 300 kHz	10 km - 1 km	LF (Low Frequency/Long Wave)
300 kHz - 3 MHz	1 km - 100 m	MF (Medium Frequency/Medium Wave)
3MHz - 30 MHz	100 m - 10 m	HF (High Frequency/Short Wave)
30 MHz - 300 MHz	10 m - 1 m	VHF (Very High Frequency)
300 MHz - 3 GHz	1 m - 10 cm	UHF (Ultra High Frequency)
3 GHz - 30 GHz	10 cm - 1 cm	SHF (Super High Frequency)
30 GHz - 300 GHz	1 cm - 1 mm	EHF (Extremely High Frequency)

Table 3.4.1. The radio wave bands (after Tascione 1988).

Frequency parameters

The propagation of radio waves in the ionosphere is influenced by certain natural frequencies that characterize the medium itself. The most important of these are the *electron gyro frequency*, f_{ge} , the *ion gyro frequency*, f_{gi} , the *collision frequency of the electrons*, ν_e , and the *plasma frequency*, f_{pe} .

The collision frequency of the electrons is simply the average number of times per second a given electron collides with another particle. The other frequencies are defined below. In the expressions given have also been included notations for the corresponding angular frequencies (denoted by the symbol ω).

$$f_{ge} = \frac{1}{2\pi} \omega_{ge} = \frac{1}{2\pi} \frac{eB}{m_e} \quad (3.4.1)$$

$$f_{gi} = \frac{1}{2\pi} \omega_{gi} = \frac{1}{2\pi} \frac{q_i B}{m_i} \quad (3.4.2)$$

$$f_{pe} = \frac{1}{2\pi} \omega_{pe} = \frac{1}{2\pi} \sqrt{\frac{n_e e^2}{\epsilon_0 m_e}} \quad (3.4.3)$$

Three more, often used, frequency parameters are the *ion plasma frequency*, f_{pi} , the *lower hybrid frequency*, f_{lh} , and the *upper hybrid frequency*, f_{uh} , which are defined as follows

$$f_{pi} = \frac{1}{2\pi} \omega_{pi} = \frac{1}{2\pi} \sqrt{\frac{n_e q_i^2}{\epsilon_0 m_i}} \quad (3.4.4)$$

$$f_{lh} = \frac{1}{2\pi} \omega_{lh} = \frac{1}{2\pi} \sqrt{\omega_{ge} \omega_{gi}} \frac{\omega_{pe}}{\sqrt{\omega_{pe}^2 + \omega_{ge}^2}} \quad (3.4.5)$$

$$f_{uh} = \frac{1}{2\pi} \omega_{uh} = \frac{1}{2\pi} \sqrt{\omega_{pe}^2 + \omega_{ge}^2} \quad (3.4.6)$$

Collision Frequency and Damping

As shown in Fig. 3.2.2, the collision frequency of the electrons is a couple of powers of ten higher in the D layer than in the upper layers. As a consequence, electrons that are set in motion by a radio wave will lose part of the energy imparted to it instead of giving it back to the wave. Therefore the wave is damped. In the upper ionospheric layers the damping of radio waves is much weaker because of the much lower collision frequency.

Phase and Group Velocity

The *phase velocity* is the velocity with which the phase of a radio wave of a single frequency propagates.

The *group velocity* is the velocity with which a wave packet composed of waves with more than one frequency propagates. The group velocity is also the velocity with which the *energy* of the wave packet propagates and therefore the velocity at which *information* can be transmitted. The group velocity can, as shown by the theory of relativity, never exceed the velocity of light. On the other hand, the phase velocity may well do so.

A radio wave in an unmagnetized plasma has a phase velocity given by the expression

$$v_{ph} = \frac{c}{\sqrt{1 - f_{pe}^2 / f^2}} \quad (3.4.7)$$

whereas the group velocity is

$$v_{gr} = c\sqrt{1 - f_{pe}^2/f^2} \quad (3.4.8)$$

Thus the product $v_{ph} \cdot v_{gr}$ is equal to c^2 . In other words, the geometric mean value between the phase and group velocities is equal to the velocity of light.

Just as in optics one can define an *index of refraction* for radio waves, which is simply given by

$$n = c/v_{ph} \quad (3.4.9)$$

Just as in optics total reflection can occur, if a wave with an angle of incidence α , encounters a boundary to a medium with higher phase velocity (lower index of refraction), and $v_{f1}/v_{f2} < \cos \alpha$. Although there are no sharp boundaries in the ionosphere, the phase velocity can change rapidly when f_{pe}/f approaches 1. As a result, a radio wave with a frequency f , propagating at an angle α to the vertical, is reflected at the level, where the *critical frequency*, f_{cr} , given by

$$f_{cr} = \frac{9}{\cos \alpha} \sqrt{n_e} \quad (3.4.10)$$

(where n_e is measured in electrons per m^3), equals the frequency of the wave.

For the D layer, where the electron density is approximately 10^9 m^{-3} , the critical frequency for vertical incidence is about 300 kHz, whereas at the F layer maximum (with electron densities of more than 10^{12} m^{-3}) it can be between 10 and 30 MHz.

In the context of using ionosondes, the quantitative translation of the ionograms into altitude profiles of electron density requires that corrections are made for the actual, curved, propagation path of the wave and for the actual variations of electron density along it, as well as for the influence of the geomagnetic field on the propagation.

A more comprehensive treatment of radio wave propagation is given in section 3.9. Here will only be given a few simple results.

A radio wave propagating parallel to the geomagnetic field can be split into two circularly polarized components, whose electric field vectors, seen in the direction of the magnetic field, rotate in, respectively, the right-hand and left-hand sense. These two partial waves have different phase velocities, namely

$$v_L = \frac{c}{\sqrt{1 - \frac{f_{pe}^2}{f^2 + f f_{ge}}}} \quad (3.4.11a)$$

for the left-hand rotating wave, and

$$v_R = \frac{c}{\sqrt{1 - \frac{f_{pe}^2}{f^2 - f f_{ge}}}} \quad (3.4.11b)$$

for the right-hand rotating one.

As the circularly polarized partial waves, which together constitute the plane polarized wave, propagate with different velocities, their relative phase changes continually during the propagation. Together they still form a plane polarized wave, but its plane of polarization rotates as a consequence of the changing phase relation of the partial waves. This rotation of the plane of polarization is called *Faraday rotation* and can be used for measuring path-integrated electron density in the ionosphere and other space plasmas.

In the case $f_{ge} \ll f$ it is found from (3.4.11a) and (3.4.11b) that the phase change per unit length, Ω_0 , becomes

$$\Omega_0 = \frac{\pi}{c} f_{pe}^2 f_{ge} / f^2 \quad \text{radians/m} \quad (3.4.12)$$

If, instead, the radio wave propagates *transverse* to the magnetic field, it is split into an *ordinary wave*, whose electric field oscillates parallel to the magnetic field, and *extra ordinary wave*, whose electric field oscillates transverse to the magnetic field. Their respective phase velocities are

$$v_O = \frac{c}{\sqrt{1 - \frac{f_{pe}^2}{f^2}}} \quad (3.4.13a)$$

$$v_E = \frac{c}{\sqrt{1 - \frac{f_{pe}^2}{f^2} \cdot \frac{(f^2 - f_{pe}^2)}{(f^2 - f_{uh}^2)}}} \quad (3.4.13b)$$

As radio communication via satellite requires that the radio waves can penetrate the density maximum of the ionosphere, frequencies exceeding the critical frequency of the F layer maximum are required. Therefore frequencies in the UHF or SHF bands are used (cf. Table 3.4.1).

However, not even these waves are entirely uninfluenced by the ionosphere they have to pass through. The most important disturbing effect on their propagation comes from rapid spatial electron density variations along their propagation path (mostly around the F layer maximum between 225 and 400 km altitude) and it leads to so-called *scintillation*, *i.e.* rapid, usually random, variations in amplitude and/or phase. The most important cause is considered to be a kind of F layer irregularities extended parallel to the magnetic field, a phenomenon, which is called *spread-F*, and is most pronounced at low latitudes.

3.5 Radiative Ionisation in an Exponential Atmosphere

To illustrate the physics behind the origin and overall structure of the ionosphere we shall quantitatively analyze the ionization equilibrium in an exponential atmosphere under the influence of ionizing radiation, and derive the so-called *Chapman profile* (Chapman 1931). It is worth noting, however that this profile was derived and published by Pedersen (1927) a few years earlier, so that *Pedersen profile* would have been a more proper name.

Assume:

- 1) An exponential atmosphere, *i.e.* a neutral gas density given by the expression

$$n_n = n_0 e^{-z/H} \quad (3.5.1)$$

The fact that a plane isothermal atmosphere is thus distributed has already been shown in section 2.3, eq. (2.3.2).

- 2) Vertically incident ionizing radiation (at wave lengths shorter than 2400 Å, for example ultraviolet light and X-rays) with the intensity I_0 above the atmosphere.
- 3) Ionisation coefficient, a_j , recombination coefficient, a_r , and absorption coefficient, a_a , are constant.

Because of absorption the radiation intensity diminishes according to the equation

$$\frac{dI}{dz} = -I n_n a_a \quad (3.5.2)$$

The number of ionizations per unit of time and volume is

$$q = a_j I n_n \quad (3.5.3)$$

The electron density, n_e , in equilibrium is determined by the balance between ionization and recombination,

$$q = a_r n_e^2 \quad (3.5.4)$$

Combination of (3.5.1) and (3.5.2) and integration gives

$$I = \text{const.} e^{-H a_a n_0 e^{-z/H}} \quad (3.5.5)$$

where the constant of integration is determined by the condition that $I = I_0$ when $z = \infty$, *i.e.*

$$I = I_0 e^{-H a_a n_0 e^{-z/H}} \quad (3.5.6)$$

Inserting (3.5.1) and (3.5.6) into (3.5.3) and application of the condition $q = r$ leads by means of (3.5.4) to the result

$$n_e = \left\{ \frac{a_j}{a_r} n_0 e^{-z/H} I_0 e^{-H a_n n_0 e^{-z/H}} \right\}^{1/2} \quad (3.5.7)$$

By differentiation it is seen that n_e has a maximum at the level

$$z = z_{max} = H \ln(a_n n_0 H) \quad (3.5.8)$$

By counting altitude from this level, *i.e.* by the substitution

$$\xi = z - z_{max} \quad (3.5.9)$$

one finds that

$$n_e = n_{e \max} e^{\frac{1}{2}(1 - \xi/H - e^{-\xi/H})} \quad (3.5.10)$$

which is the formula for the so-called *Chapman profile*.

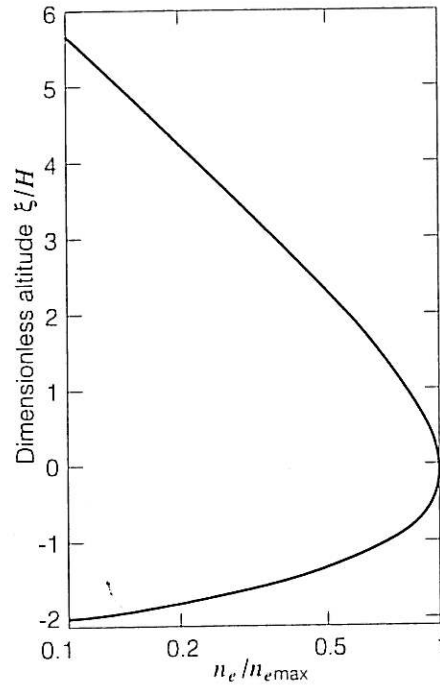


Figure 3.5.1. The Chapman profile (Hargreaves 1979).

The shape of the profile reflects the facts, already mentioned in section 3.1, that at high altitude there is a high radiation intensity but few particles to ionize. At low altitude there are many particles but little radiation left. Somewhere in between there must be a maximum of the product In_n in the equation (3.5.3) and thereby for the production of electrons.

The real ionosphere has the more complicated structure described in section 3.2, because the atmosphere is composed by several gases, which interact by complex photochemical processes, and the sun's radiation contains a variety of wavelengths.

3.6 Diffusive Equilibrium in Multicomponent Plasma

The electrically charged particles in the ionosphere are light negative electrons and much heavier positive ions, the latter of several different species. If these particles, like those in the neutral atmosphere, were to assume altitude distributions determined by the balance between thermal motion, which tries to cause upward expansion, and gravity, which holds the particles back, the light electrons would have a very different altitude distribution from the heavy ions. This would lead to a separation of positive and negative charges (the former dominating below and the latter above). However, even a tiny separation of positive and negative charges creates an electric field, which prevents further separation. This upward-directed electric field makes sure that the positive and negative charges very nearly balance each other not only totally but also locally in each volume element (quasi-neutrality).

How close this balance is can be shown by using Poisson's equation to estimate the maximum relative imbalance. For the F layer with an electron density of more than 10^{12} m^{-3} and a thickness of more than a hundred kilometers, one finds that a relative unbalance of $\Delta n/n$ would give a potential drop of $2 \cdot 10^{14} \Delta n/n$ volt. For the potential not to exceed the voltage equivalent of the particles' thermal energy, which is a fraction of a volt, the relative charge imbalance has to be less than one in 10^{15} . This illustrates the concept of *quasineutrality*, which was discussed in section 1.3.

The electric field needed to keep this balance has a potential of the order of kT/e , i.e. a fraction of a volt. With a height scale of 100 km, this implies a field strength of the order of microvolts per meter, as will be shown quantitatively below. This vertical upward electric field is called the *Rosseland Field*.

With an argument like that just given, it is possible to derive the condition $\Delta n/n < (\lambda_D/l_c)^2$ given above in section 1.3. To illustrate the interplay between gravitation, diffusion and electrostatic forces we shall give a general analysis of the equilibrium of a multicomponent plasma with an arbitrary number of ion species, numbered 1, 2, ..., N under the assumption that the ion and electron populations are each isothermal (but not necessarily at the same temperature). For simplicity we also assume that the ions are singly charged. Generalization to arbitrary charge numbers Z_{ik} and different temperatures T_{ik} for different ion species is straightforward.

The equation of pressure balance is for the electrons:

$$\frac{dp_e}{dz} = kT_e \frac{dn_e}{dz} = (-m_e g - eE)n_e \quad (3.6.1)$$

and for each of the ion species:

$$\frac{dp_{ij}}{dz} = kT_i \frac{dn_{ij}}{dz} = (-m_{ij}g + eE)n_{ij}, \quad j = 1, 2, \dots, N \quad (3.6.2)$$

Balance of charges requires:

$$n_e = \sum_{j=1}^N n_{ij} \quad (3.6.3)$$

Summation of (3.6.2) over j gives

$$kT_i \frac{d}{dz} \sum_{j=1}^N n_{ij} = (-m_{ij}g + eE) \sum_{j=1}^N n_{ij}$$

which after division with $\sum_{j=1}^N n_{ij}$ and application of (3.6.3) can be written

$$kT_i \frac{1}{n_e} \frac{dn_e}{dz} = -\bar{m}_i g + eE \quad (3.6.4)$$

where the *local mean ion mass* \bar{m}_i is

$$\bar{m}_i = \frac{\sum_{j=1}^N n_{ij} m_{ij}}{\sum_{j=1}^N n_{ij}} \quad (3.6.5)$$

Thus one has

$$kT_e \frac{1}{n_e} \frac{dn_e}{dz} = -m_e g - eE \quad (3.6.1)$$

$$kT_i \frac{1}{n_e} \frac{dn_e}{dz} = -\bar{m}_i g + eE \quad (3.6.6)$$

$$kT_i \frac{1}{n_{ij}} \frac{dn_{ij}}{dz} = -m_{ij}g + eE \quad (3.6.2)$$

From (3.6.1) and (3.6.6) the strength of the electric field can be calculated

$$E = \frac{1}{e} \frac{\bar{m}_i T_e - m_e T_i}{T_e + T_i} g \approx \frac{\bar{m}_i g}{e} \frac{T_e}{T_e + T_i} \quad (3.6.7)$$

and we can confirm that it is of the order of magnitude $1 \mu V/m$ as estimated before. Insertion of (3.6.7) into (3.6.2) gives for the density, n_{ij} , of a particular ion species, k ,

$$\frac{1}{n_{ij}} \frac{dn_{ij}}{dz} = -\frac{m_{ij}g}{kT_i} + \frac{\bar{m}_i T_e - m_e T_i}{T_e + T_i} \frac{g}{kT_i} \quad (3.6.8)$$

From the expression (3.6.8) for the altitude distribution of a particular ion species it can be seen that, if *more than one ion species* is present, the density need not vary monotonically for all the species. Ions that have a mass smaller than about half the local mean ion mass (or more precisely $m_{ij} < \bar{m}_i T_e / (T_e + T_i)$) *increases* its density upwards to the level where the mean ion mass \bar{m}_i decreased so much that $m_{ij} = \bar{m}_i T_e / (T_e + T_i)$. *Only after that* does the density of the ion species concerned decrease with altitude.

In the special case where only one ion species is present, it follows from (3.6.8) that

$$n_i = \text{const.} \cdot e^{-\int_0^z \frac{m_{ij} g dz}{k(T_e + T_i)}} = \text{const.} \cdot e^{-z/H} \quad (3.6.9)$$

where the *scale height* H is

$$H = \frac{k(T_e + T_i)}{m_i g} \quad (3.6.10)$$

In the absence of electrostatic forces ions with mass m_i and temperature T_i would have had the scale height $H_i = kT_i/m_i g$ and electrons with mass m_e and

temperature T_e the much greater scale height $H_e = kT_e/m_e g$. The result of the electrostatic forces is that both ions and electrons are forced to have a common scale height given by (3.6.10). When $T_e = T_i$,

$$H = 2H_i \ll H_e \quad (3.6.11)$$

i.e. the "compromise" scale height is much closer to the "free" scale height of the ions than to that of the electrons. In other words, the height distribution is mainly determined by the heavier species, *i.e.* the ions.

3.7 Motion of Charged Particles in Electric and Magnetic Fields

The complicated behaviour of a magnetized plasma is a result of the complex motion of the charged particles. In a neutral gas the atoms and molecules move in practically straight ballistic orbits between practically instantaneous collisions and unaffected by electric as well as magnetic fields.

In a plasma the orbits of the charged particles are influenced not only by collisions but also, and at all times, by the electric and magnetic fields that are virtually always present in the plasma.

In this section the motion of charged particles in electric and magnetic fields will be briefly described.

The gyration

Consider a particle with charge q and mass m , moving with the velocity \mathbf{v} in a homogeneous magnetic field \mathbf{B} . Its motion is illustrated in Fig. 3.7.1.

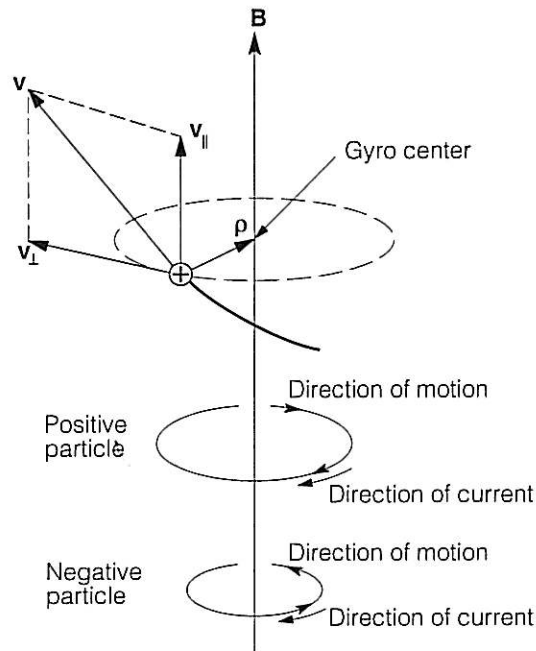


Figure 3.7.1. Unperturbed motion of an electrically charged particle in a homogeneous magnetic field.

Transverse to the magnetic field it performs a circular motion with constant angular velocity, the *gyro (angular) frequency*, ω_g , and corresponding frequency, the *gyro frequency*, f_g .

These are easily calculated from the condition that the magnetic force, $qv_{\perp}B$, must balance the centrifugal force mv_{\perp}^2/ρ . From this it follows directly the expression for the *gyro radius*

$$\varrho = \frac{mv_{\perp}}{|q|B} \quad (3.7.1)$$

and since $f_g = v_{\perp}/2\pi\varrho$, it follows that

$$f_g = \frac{1}{2\pi}\omega_g = \frac{1}{2\pi} \frac{|q|B}{m} \quad (3.7.2)$$

Seen in the direction of the magnetic field vector, *negative* particles rotate *clockwise* and *positive* particles *anticlockwise*.

In both cases the particle motion represents an electric current circulating anticlockwise around the magnetic field vector and thus generates a magnetic field, which, within the gyro circle, is directed *opposite* to the primary magnetic field (cf. Fig. 3.7.1).

If account is taken of the directions of the forces and motions, it is found that the gyro radius can be expressed as

$$\varrho = \frac{m\mathbf{v} \times \mathbf{B}}{qB^2} \quad (3.7.3)$$

where the vector ϱ represents the distance and direction from the instantaneous position of the particle to the center of the gyro circle (cf. Fig. 3.7.1)..

The magnetic moment

The circular current has a *magnetic moment* directed opposite to the primary magnetic field and is therefore called *diamagnetic* moment, whose magnitude is the product of the area of the gyro circle ($\pi\varrho^2$) and the average current (qf_g), carried by the gyrating particle. Thus the diamagnetic moment, μ , is given by

$$\mu = mv_{\perp}^2/2B \quad (3.7.4)$$

Along the magnetic field the particle moves with the constant velocity v_{\parallel} . Therefore the resulting motion becomes a spiral as illustrated in Fig. 3.7.1. The center of the circle, the *gyro center* moves at constant speed along the axis of the spiral.

Influence of an electric field

Consider next the influence of an electric field, \mathbf{E} .

Its component along the magnetic field has the only consequence that the motion becomes accelerated or decelerated.

The electric field component transverse to the magnetic field, \mathbf{E}_{\perp} , has a very different effect, namely that the particle drifts transverse to *both* the electric and magnetic field.

The simple explanation is that because of the electric field, the gyrating particle will have a somewhat greater energy, and hence larger gyro radius at one side of

the circle than the other. Therefore the circle does not close, and the particle drifts sideways as illustrated in Fig. 3.7.2.

The magnitude of the drift velocity *transverse* to the magnetic field is

$$u_{\perp} = \mathbf{E}_{\perp} / B \quad (3.7.9)$$

or in vector form,

$$\mathbf{u}_{\perp} = \mathbf{E} \times \mathbf{B} / B^2 \quad (3.7.10)$$

A straightforward way to prove this is the following. The equation of motion in the combined electric and magnetic fields is

$$m \frac{d\mathbf{v}}{dt} = q(\mathbf{v} \times \mathbf{B}) + q\mathbf{E} \quad (3.7.11)$$

Split the velocity vector \mathbf{v}_{\perp} (transverse to \mathbf{B}) into a sum of an initially unknown average velocity, \mathbf{u}_{\perp} and a superposed periodic velocity \mathbf{v}'_{\perp} , *i.e.*

$$\mathbf{v}_{\perp} = \mathbf{u}_{\perp} + \mathbf{v}'_{\perp} \quad (3.7.12)$$

By inserting (3.7.12) into (3.7.11) one finds

$$m \frac{d\mathbf{v}'_{\perp}}{dt} = q(\mathbf{v}' \times \mathbf{B}) + q\mathbf{E}_{\perp} + q\mathbf{u}_{\perp} \times \mathbf{B} - m \frac{d\mathbf{u}_{\perp}}{dt} \quad (3.7.13)$$

If the electric and magnetic fields are assumed constant in time and space, the last term in (3.7.13) is zero, and we obtain

$$m \left(\frac{d\mathbf{v}'_{\perp}}{dt} \right) = q(\mathbf{v}' \times \mathbf{B}) + q\mathbf{E}_{\perp} + q\mathbf{u}_{\perp} \times \mathbf{B} \quad (3.7.14)$$

If we now form the time average of this equation, terms containing \mathbf{v}' disappear, because \mathbf{v}' is periodic, and hence has the average zero. Then there only remains

$$0 = q\mathbf{E}_{\perp} + q\mathbf{u}_{\perp} \times \mathbf{B} \quad (3.7.15)$$

If we divide by q , and the equation is vector multiplied by \mathbf{B} , we find by means of the vector identity $\mathbf{a} \times (\mathbf{b} \times \mathbf{c}) \equiv \mathbf{b}(\mathbf{a} \cdot \mathbf{c}) - \mathbf{c}(\mathbf{a} \cdot \mathbf{b})$ that

$$0 = \mathbf{B} \times \mathbf{E} + \mathbf{B} \times (\mathbf{u}_{\perp} \times \mathbf{B}) = -\mathbf{E} \times \mathbf{B} + \mathbf{u}_{\perp} B^2 - \mathbf{B}(\mathbf{u}_{\perp} \cdot \mathbf{B}) \quad (3.7.16)$$

As $\mathbf{u}_{\perp} \cdot \mathbf{B} \equiv 0$ division by B^2 yields the expression (3.7.10) for the drift motion *transverse* to the magnetic field.

Along the magnetic field the particle is not affected by any magnetic force. If the *electric* field, \mathbf{E} , has a component, E_{\parallel} , along the magnetic field, the particle performs an accelerated motion in this direction, and the acceleration is given by

$$\frac{dv_{\parallel}}{dt} = \frac{q}{m} E_{\parallel} \quad (3.7.17)$$

Influence of a non-electric force

If the charge of the particle is q , a mechanical force, \mathbf{F} , for example gravitation or an inertial force, has the same effect as an electric force given by $\mathbf{E} = \mathbf{F}/q$.

It produces corresponding drift motions, which are illustrated in Fig 3.7.2, and which can be calculated by replacing \mathbf{E} by \mathbf{F}/q in (3.7.10) and (3.7.17), *i.e.*

$$\mathbf{u}_\perp = \mathbf{F} \times \mathbf{B} / qB^2 \quad (3.7.18)$$

and

$$\frac{dv_\parallel}{dt} = \frac{1}{m} F_\parallel \quad (3.7.19)$$

Note that the transverse drift due to an electric field is in the *same* direction for positive and negative particles, but the *opposite* if the drift is due to a mechanical force, \mathbf{F} .

Influence of a magnetic field gradient

If the magnetic field strength has a gradient transverse to \mathbf{B} , a drift motion results, because according to (3.7.1) the gyro radius is smaller where the magnetic field is strong and larger where the magnetic field strength is small, as illustrated in Fig. 3.7.2.

A quantitative expression for the drift motion due to a magnetic field gradient can be derived by noting that the gyrating particle, because of its diamagnetic moment, μ (3.7.4), is influenced by a magnetic force, which in average over the gyro orbit is given by

$$\mathbf{F}_m = -\mu \text{grad } B \quad (3.7.20)$$

Thus this average force is the same as the force from an electric field \mathbf{F}_m/q . Therefore it follows, in analogy with (3.7.10) and (3.7.17) that the transverse drift is given by

$$\mathbf{u}_\perp = \left(-\frac{\mu}{q} \text{grad } B \right) \times \frac{\mathbf{B}}{B^2} = \frac{\mu}{qB^2} \mathbf{B} \times \text{grad } B \quad (3.7.21)$$

and the motion along the magnetic field by

$$\left(\frac{d\mathbf{v}}{dt} \right)_\parallel = -\frac{\mu}{m} (\text{grad } B)_\parallel \quad (3.7.22)$$


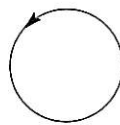
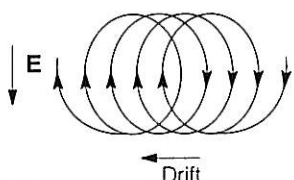
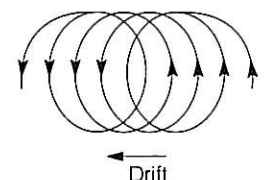
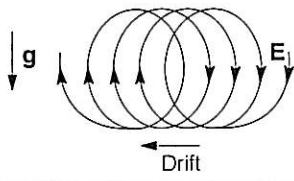
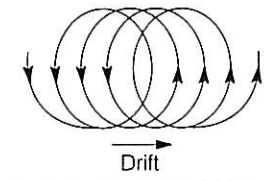
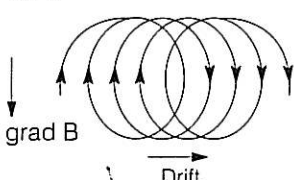
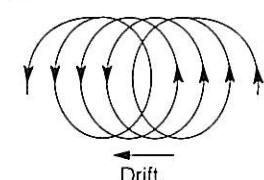
	Positive particles	Negative particles
Homogeneous magnetic field No disturbing force	$\odot B$ 	$\odot B$ 
Homogeneous magnetic field Homogeneous electric field	$\odot B$ 	$\odot B$ 
Homogeneous magnetic field Gravitation	$\odot B$ 	$\odot B$ 
Inhomogeneous magnetic field	$\odot B$ 	$\odot B$ 

Figure 3.7.2. Drift motions of electrically charged particles in a homogeneous magnetic field, under the influence of forces or gradients transverse to the magnetic field (after Alfvén and Fälthammar, 1963).

3.8 Electrical Conductivity in Magnetized Plasma

The charged particles that constitute the key ingredients of any plasma interact mutually by electrostatic forces (Coulomb forces), which have a long range, usually much larger than the average inter-particle distance, so that at every instant each particle is influenced by a large number of other particles. Furthermore, the charged particles, by charge separation, create secondary electric fields, by means of which they execute a collective interaction of a much more efficient, but also much more complicated, kind than binary collisions.

A detailed discussion of these complex processes will not be given here. We shall only note, that the description of the simultaneous many-body interaction of the charged particles can formally be reduced to an equivalent binary formalism by introducing an equivalent “collision cross section for distant encounters”, and corresponding “collision frequencies”, which allow treating Coulomb collisions as if they were binary.

3.8.1 Macroscopic Equations for Magnetized Plasma

A *macroscopic* description is characterized by relating the motion of all particles (of a certain kind) in a volume element in the plasma to the forces acting on these particles. For a strongly collision-dominated plasma one can sometimes use a *one-fluid model*, where all the particles in a volume element are assumed to move together with a common (mean) velocity. For thinner plasma, such as that in the ionosphere, and many other parts of space, it is appropriate to use a description, where neutral gas, electron gas, and (one or more) ion gas(es) constitute fluids, which move relative to each other, *i.e.* a *three(or more)-fluid model*. In one and the same volume element, electrons, ions, and neutral-gas particles can have different velocities, \mathbf{v}_e , \mathbf{v}_i , och \mathbf{v}_n .

Quasineutrality requires that the densities of ions and electrons, n_e and n_i , are (essentially) equal (we assume here singly charged ions) and this common density can be denoted n . The equations of motion for the electron, ion and neutral gases are

$$nm_e \frac{d\mathbf{v}_e}{dt} = -ne[\mathbf{E} + \mathbf{v}_e \times \mathbf{B}] - \text{grad } p_e - nm_e \nu_{en}(\mathbf{v}_e - \mathbf{v}_n) - nm_e \nu_{ei}(\mathbf{v}_e - \mathbf{v}_i) \quad (3.8.1)$$

$$nm_i \frac{d\mathbf{v}_i}{dt} = ne[\mathbf{E} + \mathbf{v}_i \times \mathbf{B}] - \text{grad } p_i - nm_i \nu_{in}(\mathbf{v}_i - \mathbf{v}_n) + nm_e \nu_{ei}(\mathbf{v}_e - \mathbf{v}_i) \quad (3.8.2)$$

$$n_n m_n \frac{d\mathbf{v}_n}{dt} = -\text{grad } p_n + nm_e \nu_{en}(\mathbf{v}_e - \mathbf{v}_n) + nm_i \nu_{in}(\mathbf{v}_i - \mathbf{v}_n) \quad (3.8.3)$$

where ν_{en} , ν_{in} and ν_{ei} are the “effective” collision frequencies for exchange of momentum by, respectively, electron–neutral, ion–neutral– and electron–ion collisions.

Here it has, for simplicity, been assumed only one species of ions and one species of neutral gas. Generalization in this respect is straightforward.

3.8.2 The Generalized Ohm's Law

When \mathbf{v}_e and \mathbf{v}_i are nearly equal, which is often the case, (even when strong electric currents flow), it is often useful to replace them with their weighted average velocity, *i.e.* the joint mass velocity of the ion- and electron gases,

$$\mathbf{v} = \frac{m_i \mathbf{v}_i + m_e \mathbf{v}_e}{m_i + m_e} \quad (3.8.4)$$

and a measure of their differential velocity, namely the electric current density,

$$\mathbf{i} = en(\mathbf{v}_i - \mathbf{v}_e) \quad (3.8.5)$$

For substitution in (3.8.1) and (3.8.2) we can invert (3.8.4) and (3.8.5) to

$$\mathbf{v}_i = \mathbf{v} + \frac{m_e}{en(m_i + m_e)} \mathbf{i} \approx \mathbf{v} + \frac{m_e}{m_i} \frac{\mathbf{i}}{en} \quad (3.8.6)$$

$$\mathbf{v}_e = \mathbf{v} - \frac{m_i}{en(m_i + m_e)} \mathbf{i} \approx \mathbf{v} - \frac{\mathbf{i}}{en} \quad (3.8.7)$$

$$\rho_m = n(m_i + m_e) \quad (3.8.8)$$

From the equations (3.8.1) – (3.8.3) one can now derive on one hand an equation for the mass velocity, \mathbf{v} , which is quite complicated and will not be discussed here, and on the other an equation for the current density \mathbf{i} .

The latter constitutes the *Generalized Ohm's Law* and can be written

$$\frac{\mathbf{i}}{\sigma} + \frac{1}{en} \mathbf{i} \times \mathbf{B} + \frac{m_e}{ne^2} \frac{\partial \mathbf{i}}{\partial t} = \mathbf{E}_{eff} \quad (3.8.9)$$

where

$$\mathbf{E}_{eff} = \mathbf{E} + \mathbf{v} \times \mathbf{B} + \frac{1}{en} (\text{grad} p_e - \frac{m_e}{m_i} \text{grad} p_i) + \beta(\mathbf{v} - \mathbf{v}_n) \quad (3.8.10)$$

and the *conductivity* σ and the friction factor β are given by the expressions

$$\sigma = \frac{ne^2}{m_e \nu_{ei} + \nu_{en} + \frac{m_e}{m_i} \nu_{in}} = \frac{ne^2}{m_e \nu_e} \quad (3.8.11)$$

where in the right hand term has been introduced a composite electron collision frequency

$$\nu_e = \nu_{ei} + \nu_{en} + \frac{m_e}{m_i} \nu_{in} \quad (3.8.12a)$$

and

$$\beta = \frac{m_e}{e} (\nu_{en} - \nu_{in}) \quad (3.8.12b)$$

The Generalized Ohm's Law is in its general form fairly complicated. But it still has one "simple" property: it still represents a *local* relation between *macroscopic* quantities. In even thinner space plasma, *e.g.* in the magnetosphere, even this local property can be lost.

In various application the general form of Ohm's law can often be simplified in ways that depend on the circumstances. If the simplification can go so far that only the first term remains on the left hand side of (3.8.9) and only the first term on the right hand side of (3.8.10), it is reduced to the ordinary Ohm's law.

3.8.3 Anisotropic Conductivity. The Conductivity Tensor.

It is clear from the form of the Generalized Ohm's Law that an electric field in one direction need not produce a current in the same direction. Because of the influence of the magnetic field, which makes it easy for electrons to move along the magnetic field but difficult to move across it, an asymmetry is present even when non-electric terms, such as pressure gradients, can be neglected.

A case that is important in the ionosphere is that where one can disregard pressure gradients and time derivatives in the Generalized Ohm's Law. One can then solve (3.8.9) for the components of the electric current density. It is suitable to choose these components in the way defined in Fig. 3.2.3, namely:

- (1) the component parallel to the component (\mathbf{E}_\perp) of the electric field that is transverse to the magnetic field, the *Pedersen current (density)*
- (2) the component transverse to both the magnetic and electric fields, the *Hall current (density)* and
- (3) the component parallel to the magnetic field, the *parallel current (density)*.

After some vector manipulations starting from the basic equations (3.8.1) – (3.8.3) one can derive the following simple relations between the components of the electric field and current densities:

$$i_P = \sigma_P E_\perp \quad (3.8.13a)$$

$$i_H = \sigma_H E_\perp \quad (3.8.13b)$$

$$i_\parallel = \sigma_\parallel E_\parallel \quad (3.8.13c)$$

where

$$\sigma_P = \sigma_e \frac{1}{1 + \omega_{ge}^2 \tau_e^2} + \sigma_i \frac{1}{1 + \omega_{gi}^2 \tau_i^2} \quad (3.8.14a)$$

$$\sigma_H = \sigma_e \frac{\omega_{ge} \tau_e}{1 + \omega_{ge}^2 \tau_e^2} - \sigma_i \frac{\omega_{gi} \tau_i}{1 + \omega_{gi}^2 \tau_i^2} \quad (3.8.14b)$$

$$\sigma_\parallel = \sigma_e + \sigma_i \quad (3.8.14c)$$

where $\sigma_e = e^2 n \tau_e / m_e$, $\sigma_i = e^2 n \tau_i / m_i$ and τ_e och τ_i are effective collision times (inverses of effective collision frequencies) for electrons and ions.

The equations (3.8.14a) – (3.8.14c) can be condensed into one single tensor equation,

$$\mathbf{i} = \sigma \cdot \mathbf{E} \quad (3.8.15)$$

where, with a suitable choice of coordinate directions, the *conductivity tensor* σ takes the simple form

$$\sigma = \begin{pmatrix} \sigma_P & \sigma_H & 0 \\ -\sigma_H & \sigma_P & 0 \\ 0 & 0 & \sigma_{\parallel} \end{pmatrix} \quad (3.8.16)$$

3.8.4 The Cowling conductivity

In some cases the Hall current is prevented from flowing, *e.g.* because the conductive region is limited to a band-like region, extended in the direction of the applied electric field, as in Fig. 3.8.1. This situation can prevail, for example, in the bands of enhanced ionization created by auroral arcs (cf. section 4.5).

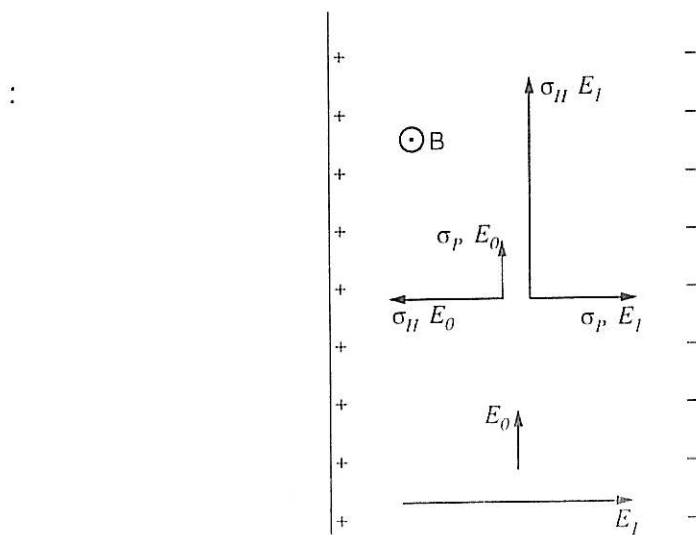


Figure 3.8.1. If the Hall current is prevented from flowing, for example by the conductive region being limited to a band extended in the direction of the applied electric field, a polarization field is formed, which leads to an enhanced effective conductivity — Cowling conductivity.

When, as in Fig. 3.8.1, the Hall current driven by the primary electric field E_0 , *i.e.* $\sigma_H E_0$, reaches the sides of the conducting band, and cannot continue, the sides of the band are charged, so that a secondary electric field, the *polarization field*, E_1 , is formed and grows, until its Pedersen current $\sigma_P E_1$ precisely balances $\sigma_H E_0$, which implies that $E_1 = (\sigma_H / \sigma_P) E_0$.

The electric field component E_1 also drives a Hall current, $\sigma_H E_1$, which adds to the Pedersen current of the primary electric field. The sum of these two currents is

$$i_{tot} = \sigma_P E_0 + \sigma_H \left(\frac{\sigma_H}{\sigma_P} E_0 \right) = \left(\sigma_P + \frac{\sigma_H^2}{\sigma_P} \right) E_0 \quad (3.8.17)$$

The resulting “effective” conductivity is called the *Cowling conductivity* and is thus given by the expression

$$\sigma_C = \sigma_P + \frac{\sigma_H^2}{\sigma_P} = \sigma_P \left(1 + \frac{\sigma_H^2}{\sigma_P^2} \right) \quad (3.8.18)$$

In the ionosphere the quotient σ_H/σ_P is often 3–4, *i.e.* σ_C about a power of ten larger than σ_P .

3.9 The Appleton–Hartree–Equation

In this section we will derive a fundamental equation for the study of the propagation of electromagnetic waves in the ionosphere. The equation, named after the English scientists Appleton and Hartree, also has important applications in the magnetosphere and in other parts of space. It is based on the following assumptions:

- 1). The motion of the ions can be disregarded. This assumption is justified at frequencies much higher than the gyro frequency because the ions are much heavier than the electrons.
- 2). The motion of the electrons is described by the equation

$$m_e \ddot{\mathbf{r}} = -e\mathbf{E} - e\dot{\mathbf{r}} \times \mathbf{B}_0 - \nu_e m_e \dot{\mathbf{r}} \quad (3.9.1)$$

where \mathbf{r} is the instantaneous position vector of an electron relative to its average position and \mathbf{B}_0 is an unperturbed homogeneous magnetic field (which is locally a good approximation for the geomagnetic field). The wave magnetic field \mathbf{b} , which is much weaker, is neglected in the equation of motion, and ν_e is the effective collision frequency of the electrons. The value of ν_e in the ionosphere is shown in Fig 3.2.2.

- 3). The electric and magnetic fields of the wave, \mathbf{E} and \mathbf{b} obey the Maxwell equations

$$\text{rot } \mathbf{h} = \frac{\partial \mathbf{D}}{\partial t} \quad (3.9.2)$$

$$\text{rot } \mathbf{E} = -\frac{\partial \mathbf{b}}{\partial t} \quad (3.9.3)$$

where

$$\mathbf{b} = \mu_0 \mathbf{h} \quad (3.9.4)$$

$$\mathbf{D} = \epsilon_0 \mathbf{E} + \mathbf{P} \quad (3.9.5)$$

and the polarization, \mathbf{P} , of the medium (electric dipole moment per unit volume), is given by the expression

$$\mathbf{P} = -ner \quad (3.9.6)$$

- 4). The waves are *plane*, *i.e.* the quantities representing the wave vary as (the real part of) the factor

$$e^{i(\omega t - kz)} \quad (3.9.7)$$

This means that the coordinate system has been oriented in such a way that the wave number vector \mathbf{k} is directed along the z -axis. Of the quantities in the exponential the *angular frequency*, $\omega = 2\pi f$, is a real number, whereas the *wave number*, k can be a complex number.

From the form of (3.9.7) there follow the simple differentiation rules:

$$\frac{\partial}{\partial t} \Rightarrow i\omega \quad (3.9.8)$$

$$\frac{\partial}{\partial z} \Rightarrow -ik \quad (3.9.9)$$

Already from the Maxwell equations (3.9.2) and (3.9.3), which in component form read,

$$ikh_y = i\omega D_x \quad (3.9.2a)$$

$$-ikh_x = i\omega D_y \quad (3.9.2b)$$

$$0 = i\omega D_z \quad (3.9.2c)$$

$$ikE_y = -i\omega b_x \quad (3.9.3a)$$

$$-ikE_x = -i\omega b_y \quad (3.9.3b)$$

$$0 = -i\omega b_z \quad (3.9.3c)$$

and the relation $\mu_0\epsilon_0 = c^{-2}$ it follows that:

a) the *index of refraction*, n , defined as

$$n = \frac{c}{v_{ph}} = \frac{c}{(\omega/k)} = \frac{ck}{\omega} \quad (3.9.10)$$

is given by the expression

$$n^2 = \frac{D_x}{\epsilon_0 E_x} = \frac{D_y}{\epsilon_0 E_y} = 1 + \frac{P_x}{\epsilon_0 E_x} = 1 + \frac{P_y}{\epsilon_0 E_y} \quad (3.9.11)$$

b) the wave is transversal in the vectors \mathbf{D} and \mathbf{b} (because $D_z = b_z = 0$).

In general it is, however, not transversal in \mathbf{E} .

c) $E_x b_x + E_y b_y + E_z b_z = 0$ i.e. $\mathbf{E} \cdot \mathbf{b} = 0$.

Note, however, that this does not mean that $\mathbf{E} \perp \mathbf{b}$. For $\mathbf{E} \perp \mathbf{b}$ it is required that $\text{Re}(\mathbf{E}) \cdot \text{Re}(\mathbf{b}) = 0$, which is in general not true.

The *polarization*, ϱ , of the *wave* is defined as

$$\varrho = \frac{E_y}{E_x} \quad (3.9.12)$$

It can easily be shown that a real-valued ϱ corresponds to linear polarization of the wave and $\varrho = \pm i$ to circular polarization of the wave. (Other values of ϱ correspond to elliptic polarization).

The *polarization*, \mathbf{P} , of the *medium (plasma)* has been defined in (3.9.6). If the differentiation rules (3.9.8) and (3.9.9) are applied to (3.9.1), and the vector \mathbf{P} is introduced by means of (3.9.6), the result is

$$m_e(i\omega)^2 \frac{\mathbf{P}}{-n_e e} = -e\mathbf{E} - ei\omega \frac{\mathbf{P}}{-n_e e} \times \mathbf{B}_0 - \nu_e m_e i\omega \frac{\mathbf{P}}{-n_e e}$$

i.e.

$$-\frac{n_e e^2}{m_e \omega^2} \mathbf{E} = (1 - i\frac{\nu_e}{\omega}) \mathbf{P} + i\frac{eB_0}{m_e \omega} \cdot \frac{\mathbf{B}_0}{B_0} \times \mathbf{P} \quad (3.9.13)$$

If now the electron plasma frequency

$$\omega_{pe} = 2\pi f_{pe} = \sqrt{\frac{n_e e^2}{\epsilon_0 m_e}} \quad (3.9.14)$$

and the electron gyro frequency

$$\omega_{ge} = 2\pi f_{ge} = \frac{eB_0}{m_e} \quad (3.9.15)$$

are introduced, (3.9.13) can be written in the simple form

$$-\epsilon_0 \left(\frac{\omega_{pe}}{\omega} \right)^2 \mathbf{E} = (1 - i\frac{\nu_e}{\omega}) \mathbf{P} + i\frac{\omega_{ge}}{\omega} \hat{\mathbf{B}}_0 \times \mathbf{P} \quad (3.9.16)$$

where $\hat{\mathbf{B}}_0$ is a unit vector in the direction of the primary magnetic field. This equation gives the relation between \mathbf{E} and \mathbf{P} and thereby (by virtue of (3.9.4)) between \mathbf{E} and \mathbf{D} . By solving (3.9.16) with respect to the components of \mathbf{P} and using (3.9.15) one can calculate the components of the *dielectric tensor* ϵ , which is defined by

$$\mathbf{D} = \epsilon \mathbf{E} \quad (3.9.17)$$

The explicit expression for ϵ is, however, not needed here.

If we introduce the notations standardized by URSI (Union Radio-Scientifique Internationale), namely

$$X = \left(\frac{\omega_{pe}}{\omega} \right)^2 \quad (3.9.18)$$

$$Y = -\frac{\omega_{ge}}{\omega} \hat{\mathbf{B}}_0 \quad (3.9.19)$$

$$Z = \frac{\nu_e}{\omega} \quad (3.9.20)$$

$$U = 1 - iZ \\ = 1 - i\frac{\nu_e}{\omega} \quad (3.9.21)$$

(3.9.16) takes the even simpler form

$$-\epsilon_0 X \mathbf{E} = \mathbf{P}U - i\mathbf{Y} \times \mathbf{P} \quad (3.9.22)$$

Since until now we have defined only one of the coordinate axes (namely z parallel to \mathbf{k}), we can, without loss of generality, further choose \mathbf{Y} parallel to the $x - z$ -plane.

The longitudinal and transverse components of \mathbf{Y} are Y_L and Y_T . Then the components of (3.9.22) become

$$-\epsilon_0 X E_x = U P_x + i Y_L P_y \quad (3.9.23a)$$

$$-\epsilon_0 X E_y = -i Y_L P_x + U P_y + i Y_T P_z \quad (3.9.23b)$$

$$-\epsilon_0 X E_z = -i Y_T P_y + U P_z \quad (3.9.23c)$$

As, according to (3.9.2c) and (3.9.5), $E_z = -P_z/\epsilon_0$, (3.9.23c) gives ,

$$P_z = \frac{i Y_T}{U - X} P_y \quad (3.9.24)$$

If P_z is substituted using (3.9.24) and the wave polarization ϱ using (3.9.12), (3.9.23 a) and (3.9.23 b) become

$$-\epsilon_0 X E_x = (U + \varrho i Y_L) P_x \quad (3.9.25a)$$

$$-\epsilon_0 X E_y = \{-i Y_L + \varrho U - \varrho Y_T^2/(U - X)\} P_x \quad (3.9.25b)$$

which combine to

$$\varrho^2 - i \varrho Y_T^2 / \{(U - X) Y_L\} + 1 = 0 \quad (3.9.26)$$

The equation for the *wave polarisation*, (3.9.26), shows that *in general two different polarizations are possible*.

The *index of refraction*, n , is, according to (3.9.11) and (3.9.25a) given by the expression

$$n^2 = 1 - \frac{X}{U + i \varrho Y_L} \quad (3.9.27)$$

Except in the case $Y_L = 0$ the two values of the wave polarization correspond to different values of the refraction index n . This means that the magnetized plasma is *birefringent*.

For waves propagating *parallel* to the magnetic field $Y_T = 0$ and (3.9.26) give $\varrho = \pm i$ which means two oppositely *circularly polarised* wave components.

If the roots of (3.9.26) are inserted into (3.9.27), the explicit form of the Appleton-Hartree-equation is obtained:

$$n^2 = 1 - \frac{X(U - X)}{U(U - X) - \frac{1}{2} Y_T^2 \pm \left\{ \frac{1}{4} Y_T^4 + Y_L^2 (U - X)^2 \right\}^{\frac{1}{2}}} \quad (3.9.28)$$

The expression thus derived for the index of refraction describes the propagation of plane electromagnetic waves. From this can be derived for example the *phase velocity* and *damping* as follows. By means of the definition $n = ck/\omega$, (3.9.10), the exponential factor (3.9.7) can be written:

$$\begin{aligned} e^{i(\omega t - kz)} &= e^{i\omega(t - z \operatorname{Re}(k)/\omega)} e^{z \operatorname{Im}(k)} \\ &= e^{i\omega(t - \frac{z}{c} \operatorname{Re}(n))} e^{\frac{z\omega}{c} \operatorname{Im}(n)} \\ &= e^{i\omega(t - z/v_{ph})} e^{-z/z_0} \end{aligned} \quad (3.9.29)$$

where

$$v_{ph} = \frac{c}{\operatorname{Re}(n)} = \frac{\omega}{\operatorname{Re}(k)} \quad (3.9.30)$$

$$z_0 = -\frac{c}{\omega \operatorname{Im}(n)} = -\frac{1}{\operatorname{Im}(k)} \quad (3.9.31)$$

Phase velocity and damping are thus related to, respectively, the real and imaginary parts of the index of refraction.

By considering a wave packet of superposed waves of the form (3.9.7) in a narrow band of frequency and wave number, it can be shown that the group velocity, v_{gr} , is given by the expression

$$v_{gr} = \frac{d\omega}{d(\operatorname{Re}(k))} \quad (3.9.32)$$

which can also be written

$$v_{gr} = \frac{1}{\frac{d}{d\omega} \left(\frac{\omega}{v_{ph}} \right)} \quad (3.9.32a)$$

For a wave packet

$$\iiint A(\mathbf{k}) e^{i(\omega t - (k_x x + k_y y + k_z z))} dk_x dk_y dk_z \quad (3.9.33)$$

where

$$\mathbf{k} = k_x \hat{\mathbf{x}} + k_y \hat{\mathbf{y}} + k_z \hat{\mathbf{z}} \quad (3.9.34)$$

one has correspondingly

$$\mathbf{v}_{gr} = \hat{\mathbf{x}} \frac{\partial \omega}{\partial k_x} + \hat{\mathbf{y}} \frac{\partial \omega}{\partial k_y} + \hat{\mathbf{z}} \frac{\partial \omega}{\partial k_z} \quad (3.9.35)$$

which is usually written symbolically as

$$\mathbf{v}_{gr} = \frac{d\omega}{d\mathbf{k}} \quad (3.9.36)$$

From the general Appleton–Hartree–equation (3.9.28) it is easy to derive the special results mentioned in section 3.4.

Example 1: **Unmagnetized** ($\omega_{ge} = 0$) **and collisionless** ($\nu_e = 0$) **plasma.**

For a wave with the angular frequency $\omega \gg \omega_{pe}$ it follows from (3.9.28), (3.9.30) and (3.9.31) that

$$v_{ph} = \frac{c}{\sqrt{1 - (\omega_{pe}/\omega)^2}} = \frac{c}{\sqrt{1 - (f_{pe}/f)^2}} \quad (3.9.37)$$

$$z_0 = \infty \quad (3.9.38)$$

Furthermore,

$$\begin{aligned} v_{gr} &= \frac{d\omega}{dk} \\ &= \frac{1}{\frac{d}{d\omega} \left(\frac{\omega}{v_{ph}} \right)} \\ &= c \sqrt{1 - (\omega_{pe}/\omega)^2} \\ &= c \sqrt{1 - (f_{pe}/f)^2} \end{aligned} \quad (3.9.39)$$

As before we can note that the *phase velocity* is greater than the velocity of light, the *group velocity* less than the velocity of light and their geometric mean equal to the velocity of light:

$$\sqrt{v_{gr} v_{ph}} = c \quad (3.9.40)$$

Example 2: **Longitudinal propagation** ($Y_T = 0$). **In the case**

$\omega_{ge}/\omega \ll 1$ (i.e. $|Y_L| \ll 1$), $\nu/\omega \ll 1$ (i.e. $Z \approx 0$), **and** $\omega > \omega_{pl}$ (i.e. $X < 1$)

Serial expansion retaining only the first order terms: ω_g/ω and ν_e/ω yields by means of (3.9.28), (3.9.30) and (3.9.31)

$$n = \sqrt{1 - \frac{\omega_{pe}^2}{\omega^2 \pm \omega \omega_{ge}}} \quad (3.9.41)$$

in agreement with (3.4.11a) and (3.4.11b).

The plus and minus signs, respectively, correspond to the two different wave polarizations, which according to (3.4.26) are $\pm i$ (circularly polarized waves). From the difference between the corresponding phase velocities, for example, the above-mentioned *Faraday rotation* can be calculated.

If, instead, we consider transverse propagation ($Y_L = 0$) the special cases (3.4.13a) and (3.4.13b) follow from (3.9.28).

Example 3: Whistler waves

A particularly interesting mode of propagation is the so-called *whistler mode* mentioned in section 4.1.

Consider longitudinally propagating waves, *i.e.* $Y_T = 0$, in a plasma where the collision frequency, ν_e , is very small, so that $U = 1$ is very well satisfied. Then (3.9.27) takes the simple form

$$n^2 = 1 - \frac{X}{1 + i\varrho Y} \quad (3.9.42)$$

According to (3.9.26) the condition $Y_T = 0$ has as a consequence that the wave polarization ϱ has the values $\pm i$. If the frequency, ω , of the wave is much less than the gyro frequency, ω_{pe} , it follows that

$$1 \ll Y \ll X \quad (3.9.43)$$

Substitution into (3.9.42) shows that it is only the polarization $\varrho = +i$ that gives a positive value to n^2 . ($\varrho = -i$ gives $n^2 < 0$, which means that n is imaginary and the wave cannot propagate). With $\varrho = +i$ (3.9.42) becomes

$$n^2 = 1 + \frac{(\omega_{pe}/\omega)^2}{\omega_{ge}/\omega - 1} \approx \frac{\omega_{pe}^2}{\omega(\omega_{ge} - \omega)} \quad (3.9.44)$$

As, according to (3.9.10) $n = ck/\omega$, the following expressions are obtained for phase and group velocity.

Phase velocity:

$$v_{ph} = \frac{\omega}{k} = \frac{c}{n} = \frac{c\sqrt{\omega(\omega_{ge} - \omega)}}{\omega_{pe}} \quad (3.9.45)$$

Group velocity:

$$v_{gr} = \frac{d\omega}{dk} = \frac{1}{\frac{dk}{d\omega}} = \frac{c}{\frac{d}{d\omega}(\omega n)} = \frac{2c\sqrt{\omega(\omega_{ge} - \omega)^3}}{\omega_{pe} \omega_{ge}} \quad (3.9.46)$$

Thus the travel time of the signal becomes

$$T(\omega) = \frac{1}{2c} \int \frac{\omega_{pe}\omega_{ge}}{\sqrt{\omega(\omega_{ge} - \omega)^3}} ds \quad (3.9.47)$$

or, expressed in terms of frequencies instead of angular frequencies,

$$T(f) = \frac{1}{2c} \int \frac{f_{pe}f_{ge}}{\sqrt{f(f_{ge} - f)^3}} ds \quad (3.9.48)$$

where the integral is to be evaluated along the path of the signal.

4. The Magnetosphere

4.1 The Geomagnetic Field

The Near Field

The *geomagnetic field* derives from electric current systems in the interior of the Earth. At the Earth's surface it is approximately a *dipole field*. The field lines of the dipole field (*i.e.* lines which are everywhere parallel to the local magnetic field vector) has the simple form shown in Fig. 1.2a. The field is rotationally symmetric around the dipole axis.

The axis of the Earth's dipole field does not coincide with the rotational axis but forms an angle of about 11° with it. If the geomagnetic field is approximated by a dipole at the center of the Earth its axis cuts the Earth's surface in two antipodal points situated at 78.3°N , 69.0°W , respectively 78.3°S , 111.0°E . The geomagnetic field, is, however, not exactly dipolar, and a better approximation is an *eccentric dipole* which is displaced 342 km from the Earth's center in the direction 6.5°N , 161.8°E and with an axis which cuts the Earth's surface at the points 80.1°N , 82.7°W , resp 76.3°S , 121.2°E .

The poles on the Earth's surface of the *real* geomagnetic field, *i.e.* the points on the Earth's surface where the magnetic field is precisely vertical – the *dip-poles* – do not coincide with any of the poles of the dipole approximation but are located at 75°N , 101°W , resp. 67°S , 143°E .

The *strength* of the geomagnetic fields is $62\ \mu\text{T}$ (0.62 Gauss) at the poles and $31\ \mu\text{T}$ (0.31 Gauss) at the magnetic equator.

The *direction* of the geomagnetic field is such that the magnetic pole located in the Northern hemisphere is a magnetic south pole and vice versa.

A quantitative mathematical description of the geomagnetic field is given in Section 4.6.

The Distant Field

According to a well-known law of electromagnetism, the magnetic field of a closed current system becomes more and more dipole-like with increasing distance from the current system. As the geomagnetic field already at the Earth's surface is very nearly dipolar, one would expect that this would hold more and more precisely with increasing distance from the Earth.

Since the strength of the dipole field is inversely proportional to the cube of the distance from the center, one would also expect that the geomagnetic field would relatively rapidly approach zero.

The expectation that the geomagnetic field should become increasingly dipole-like at higher altitudes was initially confirmed, when measurements were made by

means of satellites. But when the satellites reached distances of many Earth radii, there was a surprise. In the direction of the Sun the field strength started to increase again and at a distance of about ten Earth radii it changed drastically. Often the magnetic field suddenly changed from its regular direction to the opposite with about the same field strength. The location of this sudden change was initially called the *Cahill discontinuity* after its discoverer. It is now called the *magnetopause*, since we know that it constitutes the boundary between the geomagnetic field (with field lines which come from the interior of the Earth and return there) and an external magnetic field, the sources of which are electric currents in interplanetary space and the field lines of which connect to the Sun.

The discovery of the magnetopause is an example of a discovery which could only be made by measurements *in situ* and not through remote observations.

In the direction away from the Sun the geomagnetic field becomes more and more radial and extends to very large distances (cf. Fig. 1.2.1c and Section 4.2).

We now know that the reason for the drastic deformation of the geomagnetic field is electric currents which are caused by a plasma flow from the Sun, the *solar wind* which is treated in Chapter 6. The result is that the geomagnetic field is limited to a cavity in the solar wind. The region of space which is dominated by the geomagnetic field is called the *magnetosphere*.

Matter Outside the Ionosphere

The occurrence of matter even beyond the ionosphere was discovered as a consequence of a certain type of interferences in telephone lines. They consisted of whistling sounds and are nowadays called "*whistlers*". They were explained by the physicist Barkhausen (known from magnetism) who showed that they are caused by audiofrequency electromagnetic waves which are excited by lightening discharges and then propagate along geomagnetic field lines between the two hemispheres of the Earth (Fig. 4.1.1). They propagate in a special wave mode, called *whistler mode*, which has been described in Section 3.9. In this mode waves of different frequencies propagate at different velocities. From a given lightening discharge, waves in the medium frequency range arrive first and later waves with higher and lower frequencies (cf. Fig. 4.1.1). Hence the whistling sound.

Whistler waves are usually registered in a diagram with the arrival time on the horizontal axis, the frequency on the vertical axis and intensity represented by blackening, known as *sonagram*. An example is shown in Fig. 4.1.2.

The time required for propagation from one hemisphere to the other is according to (3.9.47)

$$T(f) = \frac{1}{2c} \int \frac{f_{ge} f_{pe}}{\sqrt{f(f_{ge} - f)^3}} ds \quad (4.1.1)$$

The integral is to be evaluated along a geomagnetic field line. For a given field line the function $f_g(s)$ is also given. For a given density distribution, *i.e.* a given $f_{pe}(s)$, and if T_f is plotted against frequency in the way shown in Fig. 4.1.3, the result is a curve which looks similar to a whistler sonagram. It has a minimum at a certain frequency, the *nose frequency*, f_n .

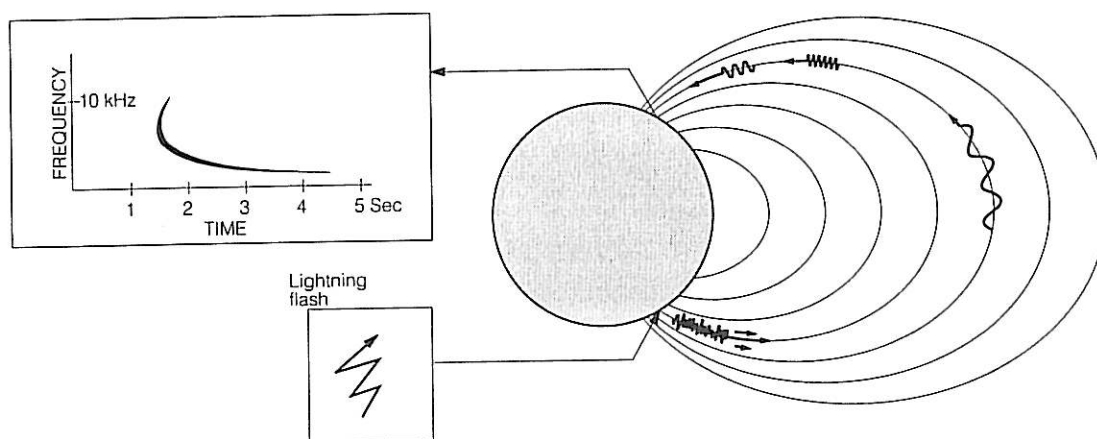


Figure 4.1.1. Audiofrequency electromagnetic waves from lightening discharges can propagate along geomagnetic field lines in a mode where low and high frequencies propagate more slowly than waves of medium frequencies.

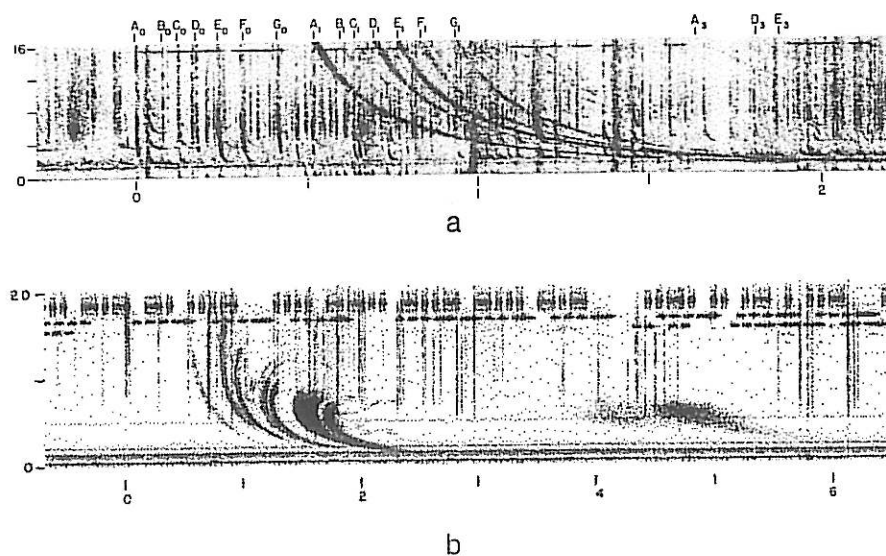


Figure 4.1.2. Example of sonagram (Helliwell 1965).

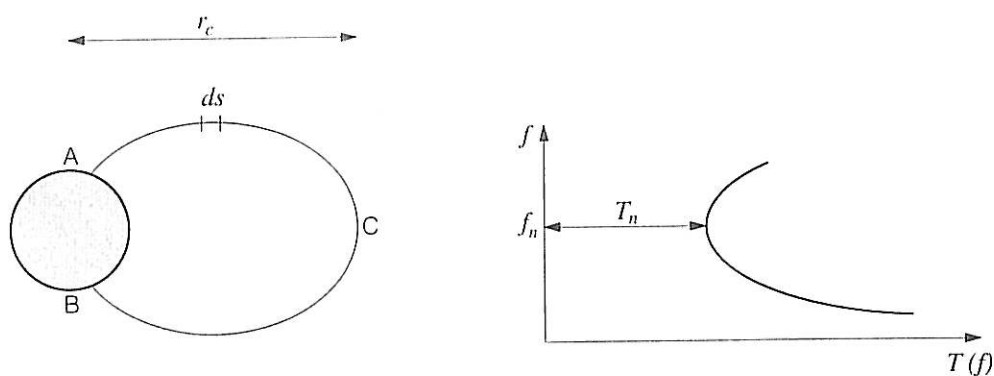


Figure 4.1.3. The path of a whistler signal in the magnetosphere and the relation between frequency and travel time from A to B

One can show that even for widely different assumptions about the density distribution the following simple relation between the nose frequency, f_n , and the electron gyro frequency at the peak of the field line (f_{ge0}) is valid :

$$f_{ge0} \approx 2.5 f_n \quad (4.1.2)$$

Thus one can directly from a sonagram read off f_n and thereby determine f_{ge0} . (This is true for “nose whistlers”, *i.e.* whistlers where the minimum falls within the registered frequency range.) Since the magnetic field is well-known (practically a dipole field) in the regions where whistlers occur, one can thereby also determine which magnetic field line the signal has followed. Thus the function $f_{ge}(s)$ is known, and one can through analysis of the sonagram determine which function $f_{pe}(s)$, *i.e.* which electron density distribution, $n_e(s)$, that gives a $T(f)$ which agrees with the observed sonagram.

Using many whistlers which have propagated along different field lines one can reconstruct the whole density distribution. An early example of such a calculated density distribution is given in Fig. 4.1.4. In this particular case the density distribution is represented by the function

$$n_e = \frac{2580}{r^3} e^{3.03/r} \quad (4.1.3)$$

(with r in Earth radii and n_e in cm^{-3}). Already in 1963 it was discovered by analysis of whistlers that the normal density distribution often terminated abruptly at a (variable) distance of between about 3 and 6 Earth radii. At this boundary, nowadays called the *plasma pause* (Section 4.2), the density drops to low values (of the order of 1 cm^{-3} or even less).

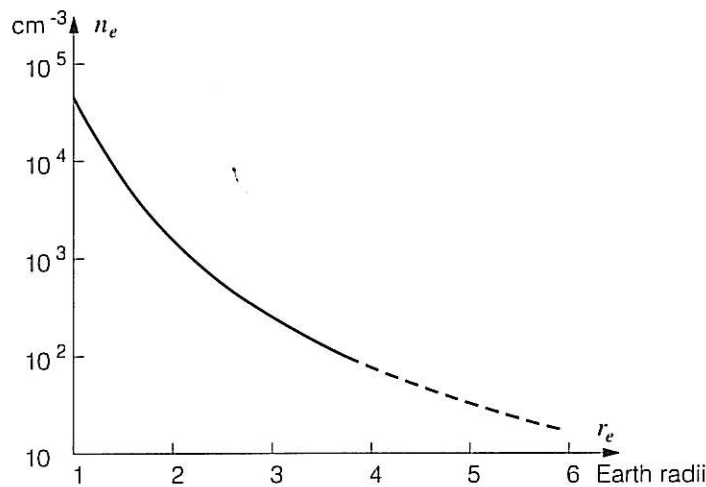


Figure 4.1.4. Example of electron density distribution in the equatorial plane derived from whistler data. (Pope 1962).

4.2 The Structure of the Magnetosphere

After many years of research, which has repeatedly led to surprising results, the structure and the material content of the magnetosphere are now essentially known.

The most important parts of the magnetosphere are illustrated in Fig.4.2.1.

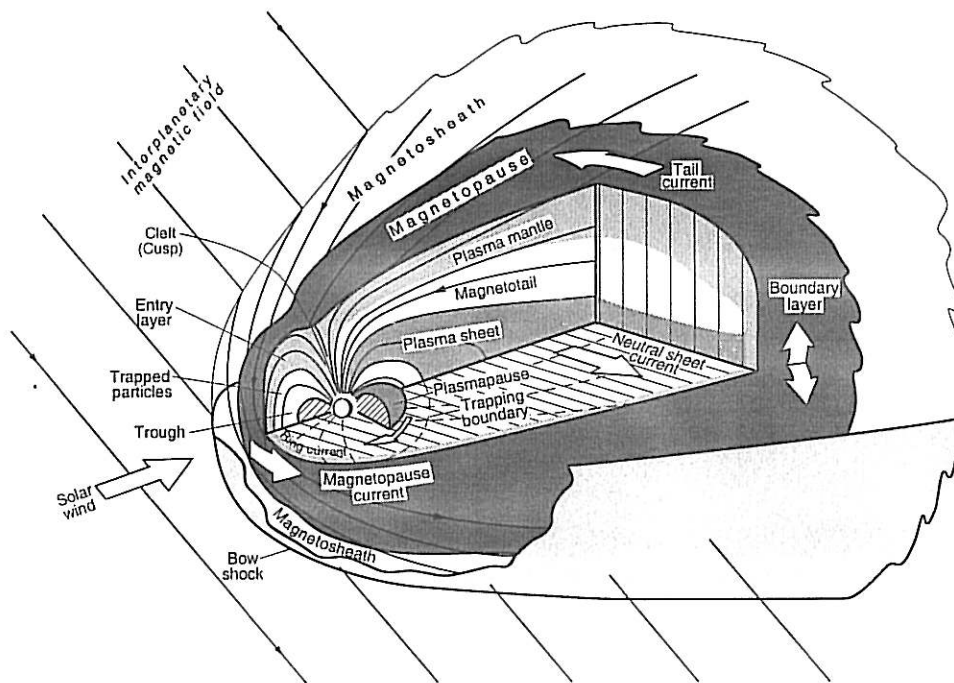


Figure 4.2.1. The Magnetosphere

The *magnetopause* is the outer boundary of the magnetosphere. On the sunward side it is located at a distance which is usually about 10 Earth radii, however, this location can vary by several Earth radii outwards and inwards depending on the strength of the solar wind and the direction of its magnetic field. Occasionally it even reaches inside the geostationary satellite orbit at about 6.6 Earth radii distance from the Earth's center, or 36 000 km above the Earth's surface, where many meteorological satellites and communication satellites are located. On the night side it extends to unknown distances of at least hundreds of Earth radii.

A couple of Earth radii ahead of the magnetopause is the *bow shock*, a shock front analogous to the bow shock ahead of a supersonic aircraft. It arises because the magnetosphere constitutes an obstacle in the supersonic flow of the solar wind, whose speed, 300 – 900 km/s, is greater than both the speed of sound (about 50 km/s) and the other relevant signal speed, the Alfvén velocity (ca 40 km/s), cf. (1.3.5).

From a physical point of view, however, the bow shock of the magnetosphere is remarkable for the following reason: All shock waves known from ordinary gas

dynamics exist thanks to the binary collisions between the molecules of the gas, and their thickness is related to the mean free path of the molecules between collisions. In the solar wind the mean free path is about 10^{12} m (a distance comparable to the distance between the Sun and the Earth!), *i.e.* the solar wind particles are practically “collisionless”. In spite of this a bow shock is created in front of the magnetosphere. The existence of collisionless shock waves had been theoretically predicted but never experimentally observed. The bow shock of the magnetosphere is the first experimental proof of the existence of the phenomenon.

Like other shock waves the bow shock of the magnetosphere represents a transition between supersonic and subsonic flow. Thus, inside the bow shock there is a region of subsonically flowing plasma. This region is called the *magnetosheath*. In the magnetosheath the plasma flows around the obstacle, the magnetosphere. Where the flow bifurcates in the middle of the front of the magnetosphere a *stagnation point* is formed.

The nightside extended magnetic field lines form the *magnetotail*. At the border between the oppositely directed magnetic field lines from the northern and southern hemispheres respectively the magnetic field changes direction. At this border surface, the *neutral sheet*, the magnetic field strength almost passes zero, but not quite, because there is a small but finite normal component of the magnetic field, which connects the northern magnetic field lines to the southern.

In the magnetosphere there are a number of well defined plasma populations.

Nearest to the Earth, on magnetic field lines originating at low and middle geomagnetic latitudes, there is a comparatively dense plasma, which forms the *plasmasphere*. This plasma is a kind of continuation of the ionosphere underneath it and from which it is formed by diffusive transport from high density region. The temperature is essentially the same as in the ionosphere, *i.e.* some thousands of K, except in the outermost regions, and the density decreases gradually from ionospheric densities to values of 10 – 100 particles per cubic centimetre.

The outer boundary of the plasmasphere, the *plasmopause*, can sometimes be very sharp (less than one Earth radius thick) but sometimes much more diffuse. During geomagnetically disturbed conditions, the outer parts of the plasmasphere are removed through a convection enforced from the outside, which causes the plasmasphere to shrink and get a sharp boundary. Subsequently the emptied regions become refilled through diffusion of new plasma from the ionosphere below, making the border again diffuse.

Outside the plasmopause, especially on the dayside, the plasma density can sink to very low values. This region of low density is usually called the *trough*.

Because of its relatively good coupling to the underlying ionosphere, the plasma sphere participates in the rotation of the Earth, it corotates. Sometimes, however, the outermost parts deviate substantially from corotation.

At higher latitudes on the nightside there is a large region, the *plasma sheet*, which is filled with a much higher energy plasma, the equivalent temperature of which is of the order of millions of degrees. The *plasma sheet boundary layer* toward higher latitudes is closely related to the processes that cause the aurora (Section 4.5 below). Magnetic field lines at latitudes higher than the plasma sheet form the *polar plumes*, of the magnetosphere, where the density of matter is extremely low, $1/100 \text{ cm}^{-3}$ or less.

Inside the front and flanks of the magnetosphere there is a region with a thickness of a couple of Earth radii, which is filled with a plasma which at least partly originates in the magnetosheath outside. This *entry layer* is a result of the still unexplained processes by which plasma from the outside penetrates the magnetopause. Inside the magnetopause on the field lines that are bent toward the nightside there is a corresponding region, discovered relatively late, which is called the *plasma mantle*.

As for the plasma that fills the magnetosphere it was assumed until recently that – except for the plasmasphere which is of ionospheric origin – it consisted of a hydrogen plasma from the solar wind and thus ultimately from the Sun. Not until the 1970's was it surprisingly discovered that during disturbed periods of time there is even on high latitudes so strong an outflow of ions from the ionosphere that large parts of the magnetosphere are dominated by plasma of terrestrial origin (from the ionosphere and thus ultimately from the atmosphere). What was particularly surprising was that this plasma often was dominated by oxygen ions, in spite of the fact that the uppermost ionosphere is dominated by helium and hydrogen. The reason for this is that the still unexplained mechanisms that cause the plasma outflow are strongly selective and represent one or more previously completely unknown separation mechanisms.

4.3 The Earth's Radiation Belts

A Scandinavian pioneer in auroral research, the Norwegian Carl Störmer, performed extensive calculations on the motions of charged particles in a magnetic dipole field (as a model of the Earth's magnetic field). His primary interest was particle orbits which from outer space lead down to the auroral zone. But he also found that there existed particle orbits, which were limited to closed regions in space. A particle in such an orbit is magnetically trapped during indefinite time within an angular region around the dipole axis (Störmer, 1955). Since the equations of motion are valid both backwards and forwards in time, it follows that it is also impossible for a particle to enter into such a region from the outside. These regions were therefore called "*forbidden regions*". It was consequently supposed that there would be no particles in such orbits.

In the 1950's F. Singer speculated on the possibility that the forbidden regions might still contain particles. Possible reasons could be perturbations of the magnetic field, collisions between particles or injection through nuclear reactions caused by cosmic radiation. These ideas were, however, not published and not paid attention to.

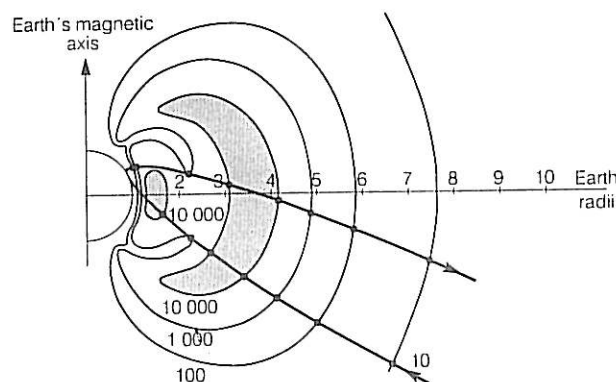


Figure 4.3.1. Intensity contours for the original Van Allen belts (Van Allen and Frank 1959).

When the American satellite Explorer I was launched, it also contained an experiment by Professor James Van Allen at the University of Iowa. It was based on a Geiger-Müller-tube for measuring cosmic radiation. As expected the observations confirmed an increase of intensity with altitude (because of the decreasing atmospheric absorption) and a theoretically predicted maximum the *Pfotzer maximum*, resulting from a high yield of secondary particles in a certain altitude range. Above that the intensity ought to approach a constant level corresponding to the flux of cosmic radiation in free space. Instead the counting rate of the Geiger-Müller-tube fell to zero. This remarkable result was found to be the consequence of the satellite being hit by so intense a particle flux that the Geiger-Müller-tube was blocked. This

was the discovery of the innermost of the Earth's radiation belts or the *Van Allen belts* (Fig. 4.3.1). These *radiation belts* constitute an important danger to satellite-borne electronics, which must be made radiation resistant and/or be surrounded by radiation shielding.

This inner radiation belt turned out to consist of protons with energies of above 30 MeV and fluxes of the order of $10^4 \text{ cm}^{-2}\text{s}^{-1}$. It was later found that these protons derived from neutrons, which had been formed by the incidence of cosmic rays against the ionosphere, ejected back into space and decayed, whereby the released proton was trapped in the magnetic field. This source came to be known as *CRAND* (*Cosmic Ray Albedo Neutron Decay*, cf. Chapter 9).

Later satellites, Explorer IV and Pioneer III, found an outer belt, which was found to consist of electrons with energies of above 1.5 MeV and very variable fluxes. The variations arise because the particles are subject to complicated transport and loss processes, which will not be dealt with in detail here (but somewhat touched on in Section 4.8). Measured fluxes of electrons above 1 MeV and protons above 10 MeV are shown in Figs. 4.3.2 and 4.3.3.

Later, when the measurements were extended to lower and lower energies, it was found that practically the whole magnetosphere except for the outermost parts, contains magnetically trapped particles, electrons as well as protons, of many different energies. At energies of a few keV and lower the particles become too sensitive to electric perturbation fields to be able to stay magnetically trapped (as will be more closely described in Section 4.8) (Van Allen and Frank 1959).

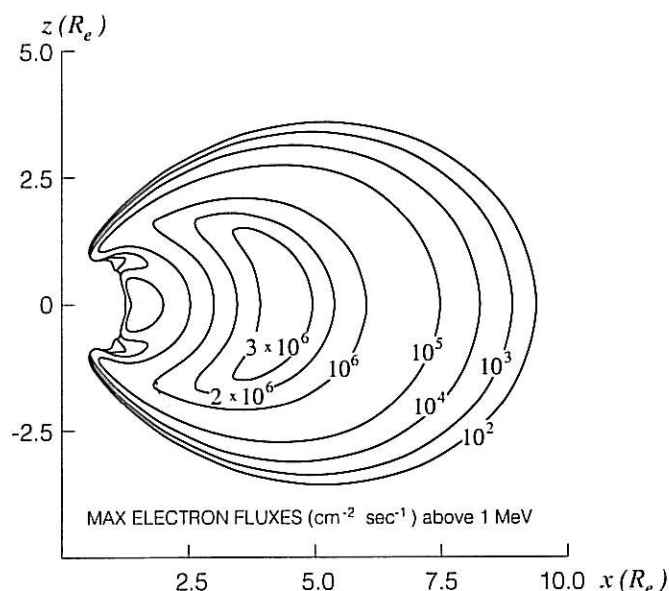


Figure 4.3.2. Intensity contours for electrons above 1 MeV (Daly 1988).

A detailed description of the radiation environment in space is found in *Handbook of Geophysics and the Space Environment* (1985). For actual calculations of radiation doses in given satellite orbits one can, however, make use of a computer program, *ESA-Base*, which is available free of charge from the *European Space Organisation*, *ESA*, via the *European Space Science and Technology Centre (ESTEC)* in Noordwijk, Holland.

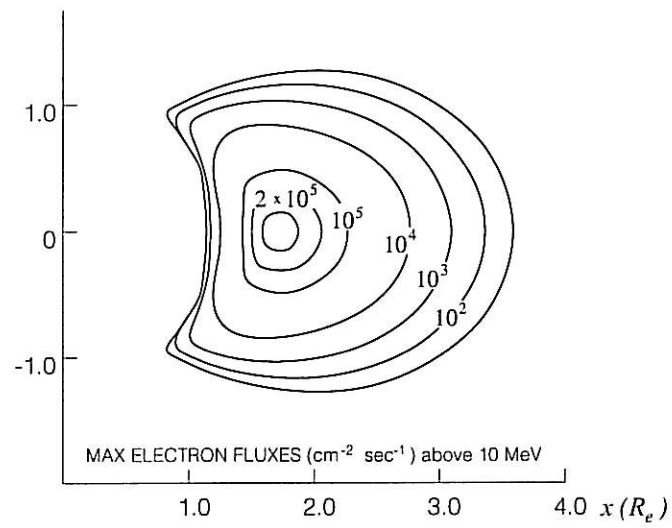


Figure 4.3.3. Intensity contours for protons above 10 MeV (Daly 1988).

4.4 Dynamics of the Magnetosphere

Pressure Balance

Since the cause of the deformation of the geomagnetic field is the solar wind, one can expect that its extent on the frontside is determined by the *dynamic pressure* of the solar wind. This is proportional to the density of the solar wind, ρ_m , and to the square of its flow velocity, v , and is given by the expression

$$p_d = \rho_m v^2 \quad (4.4.1)$$

The capability of the magnetic field to resist the pressure is proportional to the square of the magnetic field strength. For a dipole field with the magnetic moment a the field strength in the equatorial plane at radius r is given by

$$B = \left(\frac{\mu_0 a}{4\pi} \right) \frac{1}{r^3} \quad (4.4.2)$$

One can therefore expect that pressure balance, and thereby the location of the dayside border of the magnetosphere, should occur at a radius where

$$\rho_m v^2 = \left[\left(\frac{\mu_0 a}{4\pi} \right) \frac{1}{r^3} \right]^2 / 2\mu_0$$

A complete calculation taking into account also the curvature of the boundary gives the result (Axford, 1964)

$$r = \left(\frac{\mu_0 a}{4\pi} \right)^{1/3} (2\mu_0 K \rho_m v^2)^{-1/6} \quad (4.4.3)$$

where K is a numerical factor of the order of magnitude 1. This result has been found to describe reasonably well the distance to the dayside magnetopause, (the *standoff distance*).

Since the solar wind pressure, like other parameters characterizing the solar wind, varies within wide limits, so does the extent of the dayside magnetosphere. Usually the distance from the center of the Earth to the dayside magnetopause is 8 – 12 Earth radii. There are, however, also much greater variations, and several cases are known, where the magnetopause has come inside the orbit of geostationary satellites (at an altitude of 36 000 km, *i.e.* about 6.6 Earth radii from the center of the Earth).

The continual variations in the position of the magnetopause are typical for the variability of the magnetosphere in general. In fact, it is likely that the magnetosphere never assumes true equilibrium state before some new change occurs. The magnetosphere is, in other words, a strongly dynamic system.

Plasma Entry

Although the magnetopause, at least on the dayside, is a well-defined border between the geomagnetic and the interplanetary fields, plasma from the solar wind enters into magnetosphere. How the entry of the plasma takes place is not yet fully clarified, but it seems to be a matter of dynamic rather than stationary processes. For example it has been found that the electric field at the magnetopause exhibits great fluctuations, which in general completely overshadow the much smaller average value. There are reasons to believe that these fluctuations play a decisive role for plasma transport through the magnetopause, but there exists no generally accepted theory of how this occurs.

Energy Balance

The cross section area of the magnetosphere transverse to the flow direction of the solar wind is about 20 by 20 Earth radii, *i.e.* about 10^{16} m^2 . With an energy density of about $7 \cdot 10^{-10} \text{ Jm}^{-3}$ and a speed of 320 km/s the solar wind delivers a power of about $2 \cdot 10^{-4} \text{ Wm}^{-2}$, which means that over the whole cross sectional area of the magnetosphere about $2 \cdot 10^{12} \text{ W}$ is available.

What part of the solar wind energy is actually delivered to the magnetosphere is, however, very variable. During geomagnetically quiet conditions only some per cent of the available power is absorbed. In connections with magnetic storms with associated auroral activity the power fed into the magnetosphere can increase tenfold or more.

It was soon found that the power fed into the magnetosphere had little relation to the speed or the dynamic pressure of the solar wind, and thus to the available power. On the other hand it turned out to be strongly dependent on the direction of the interplanetary magnetic field and hence of the interplanetary electric field. When the interplanetary magnetic field is directed southward, *i.e.* opposite to the geomagnetic field in the equatorial plane, the energy injection becomes strong. This case corresponds to an electric field directed from the morning side of the magnetosphere to its evening side (positive E_y). When the interplanetary magnetic field is directed northward *i.e.* parallel to the geomagnetic field in the equatorial plane, the energy injection becomes weak. An example of this relation is shown in Fig. 4.4.1.

It is a well-known result from electromagnetic theory that an electric conductor moving in a magnetic field gives rise to an electric field (this is the basic principle for all electric generators). Since the solar wind is electrically conducting, it gives rise to an electric field perpendicular to both its directional motion and the magnetic field. Since it moves radially outwards from the sun, the southward interplanetary magnetic field means that the electric field is directed from the morning side to the evening side of the magnetosphere (dawn-to-dusk). The solar wind therefore acts as a magnetohydrodynamic (MHD) generator with an electromotive force of the order of 100 kV.

What has been said above about the energy release can then be formulated in such a way that the energy release is large when the electric field component in the direction from morning to evening is positive, small when it is negative. In the

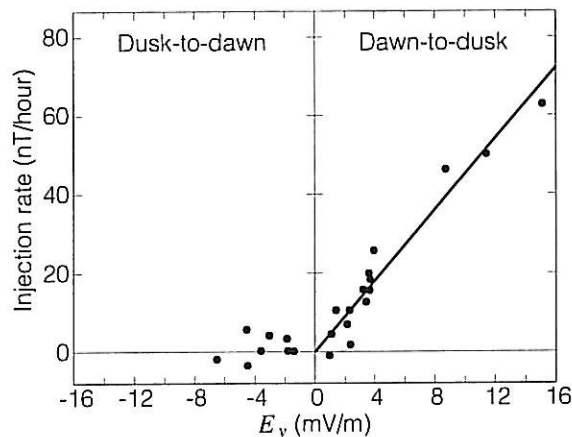


Figure 4.4.1. Power released in the magnetosphere as a function of the y -component of the interplanetary electric field (Burton *et al.* 1975). the rate of change of the geomagnetic field at the Earth is used as a measure of the energy injection rate.

light of this Fig. 4.4.1 means that the magnetosphere behaves almost precisely as if it were a simple rectifier with its “conduction direction” from morning to evening.

This is a surprising fact considering that the energy release is a result of an interplay between processes so complicated that one is still very far from being able to describe them theoretically.

Geomagnetic Disturbances

Disturbances in the geomagnetic field are strongest and most common in the auroral zone, *i.e.* the annular region around each geomagnetic pole where auroras are most frequent (cf. Section 4.5).

The disturbances are usually strongest in the horizontal component (H) and can become more than 1000 nT. (The traditional units for the geomagnetic field are 1 Gauss = 10^{-4} T, and 1 gamma = 1 nT.) Since the geomagnetic field at the Earth’s surface (apart from local anomalies) is between 0.32 and 0.62 Gauss (between 32 and 62 μ T) the change in the *total* magnetic field vector is fairly small both in terms of magnitude and direction. On the other hand, at high latitudes the horizontal component of the magnetic field is rather small (about 13 μ T at ESRANGE in Kiruna). A disturbance of the order of 1000 nT therefore represents about 8 per cent of the horizontal component in the auroral zone.

At magnetic observatories on the Earth’s surface the geomagnetic field is registered routinely. A large number of such observatories deliver their records, *magnetograms*, to a number of *World Data Centers*, which in turn deliver data on request.

Magnetospheric storms and Substorms

Periods of time when the geomagnetic field is strongly disturbed are traditionally called *geomagnetic storms*. They are associated with strong *auroral activity*. As

observed at the Earth's surface at *low* latitudes a major magnetic storm is characterized by characteristic variations in the horizontal component of the magnetic field. These variations are often characterized in terms of three phases as schematically shown in Fig. 4.4.2. During a *growth phase* with a duration of a few hours the change of the H-component is positive. The *main phase* starts with an often abrupt negative change, which can be hundreds of nT and is followed by a slow *recovery phase* with a duration of some days.

At *high* latitudes, and especially in and near the auroral oval, the variations of the magnetic field are much stronger and more irregular in both space and time due to strong local current systems.

A closer analysis of geomagnetic disturbances has led to the identification of a basic element of the geomagnetic storm, which is called *substorm*. What is observed at the Earth's surface is only a manifestation of a large scale phenomenon that really takes place in the outer magnetosphere. This phenomenon is called *magnetospheric substorm* (Akasofu, 1972).

From the point of view of the magnetosphere the most characteristic feature of the substorm is a certain type of change of the structure of the magnetic field. During a *growth phase* the geomagnetic field lines on the night side are deformed in such a way that they are pulled out to a more and more tail-like shape, which means that strong electric currents directed from the morning to the evening side of the magnetosphere (the cross-tail current) are built up, and that magnetic energy is stored in the deformed magnetic field. At the start, "*breakup*", of the substorm the shape of the magnetic field suddenly changes to a more dipole-like form. This means that the cross-tail magnetospheric currents decrease. Instead, strong electric currents occur in the ionosphere of the auroral oval and cause the decrease of the H-component observed at the Earth's surface. It is generally agreed that what happens is that the cross-tail current is "short-circuited" via the ionosphere and that the reason for this is a large-scale instability in the magnetosphere. There are, however, very divided opinions about the nature of this instability.

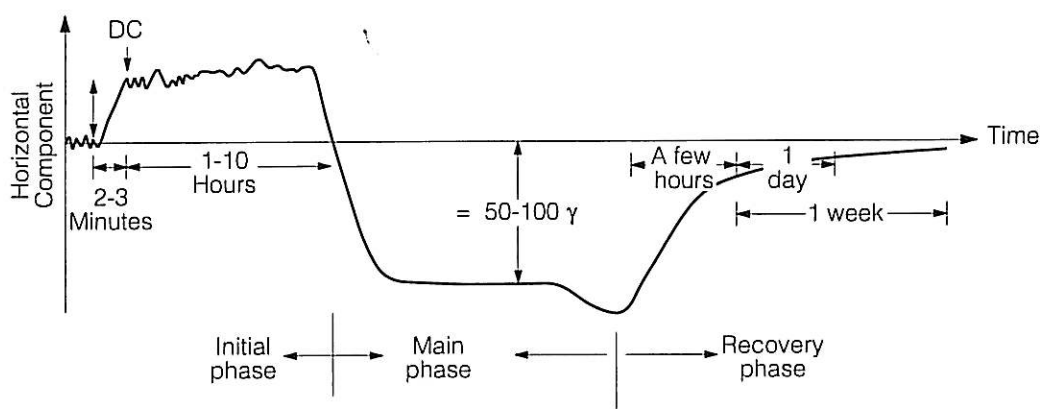


Figure 4.4.2. The variation of the horizontal component of the geomagnetic field at the Earth's surface during a magnetic storm (Tascione 1988). The unit γ ($1 \gamma = 1 \text{ nT}$) is commonly used in the context of magnetic storms.

4.5 The Aurora

The most striking expression of the dynamics of the magnetosphere is the phenomenon of the *aurora*.

Early Concepts of the Aurora

The aurora has fascinated man since time immemorial and given rise to various myths about its origin. Fig. 4.5.1 illustrates a Laplandish myth about the origin of the aurora.



Figure 4.5.1. The origin of the aurora according to a Laplandish myth (after Eather 1980).

At lower latitudes, where the aurora is rare it has sometimes caused great excitement. A famous aurora occurred in AD 79 above Italy and led to the dispatch of soldiers from Rome to Ostia, which was thought to be on fire.

The phenomenon also interested Aristotle, who was of the opinion that it arose from movements in a “fire sphere” between the Moon and the Earth. This explanation is said to have been generally accepted until the 16th century. Genuine scientific research on the aurora started with the Swedish scientist Anders Celsius, who showed that there existed a relation between the occurrence of sunspots and of auroras.

Scandinavian Pioneers

Also later, and even in our time, Scandinavians have played a prominent role in auroral research. This has a very natural explanation, because among all the areas of the Earth where the aurora is common, it is the Northern Scandinavian sector of the auroral zone that has the most favourable climate.

The foremost Scandinavian pioneers in this field of research are Kristian Birkeland (1867–1917), Carl Störmer (1874–1957) and Hannes Alfvén (1908). Kristian Birkeland made experiments in order to simulate the aurora in the laboratory. He used a magnetized sphere, "*terrella*" (diminutive form of the Latin word *terra* = the Earth) as a model of the Earth and showed that electrons from a cathode reached the *terrella* in annular regions reminiscent of the Northern and Southern auroral zones. Birkeland also predicted that the aurora had to do with electric currents flowing between the Earth and outer space. Since in those times space was considered to be a vacuum, Birkeland's theories were not taken seriously until satellite-borne measurements showed their correctness. (The same Birkeland also invented the Birkeland–Eyde-process for production of salpetre from the nitrogen of the air.)

Carl Störmer was the first to determine the altitude of the aurora by photogrammetric methods. His most important contribution was, however, very comprehensive calculations of the orbits of charged particles in magnetic fields – a task which before the era of computers was extremely laborious.

The Swedish scientist Hannes Alfvén has made outstanding pioneer contributions both within auroral research and within many other areas of cosmical physics. One such contribution was the development of the perturbation theory which provided an extreme simplification of some of the most complicated calculations of particle orbits in electric and magnetic fields. (This theory is discussed in Section 4.8.) Another pioneer contribution by Alfvén was the discovery of an entirely new kind of waves, the magnetohydrodynamic waves (Section 1.3), nowadays called Alfvén waves. For some of his achievements he was awarded the 1970 Nobel prize for physics.

Auroral Forms

The aurora takes many different forms. Some of them are illustrated in Fig. 4.5.2. The most important auroral forms are *homogeneous arcs* (with a sharp lower border and a more diffuse upper limit), *arcs with ray structure* (are often formed when a homogenous arc increases in intensity), *curtains* (curtain-like wavy belts of light, sometimes with ray structure), *rays* (luminous structures with a great altitude variation parallel to the local direction of the geomagnetic field, often rapidly disappearing and reappearing). In the evening and early night there are most often quiet, nearly motionless, auroral arcs which extend hundreds of kilometers in the east–west direction. Later in the night there are often auroral arcs that can exhibit rapid changes and motions. On the morning side there are usually diffuse and pulsating auroral forms.

When aurora occurs in magnetic zenith (*i.e.* around the geomagnetic field line through the position of the observer) it forms a *corona* (Fig. 4.5.2).

The colour of the aurora is usually yellow–green because of the dominating spectral line 5577 Å (from oxygen atoms). Sometimes red colours occur. (Weak

auroras look white to the eye because the most light-sensitive sensors in the retina cannot distinguish colours.)

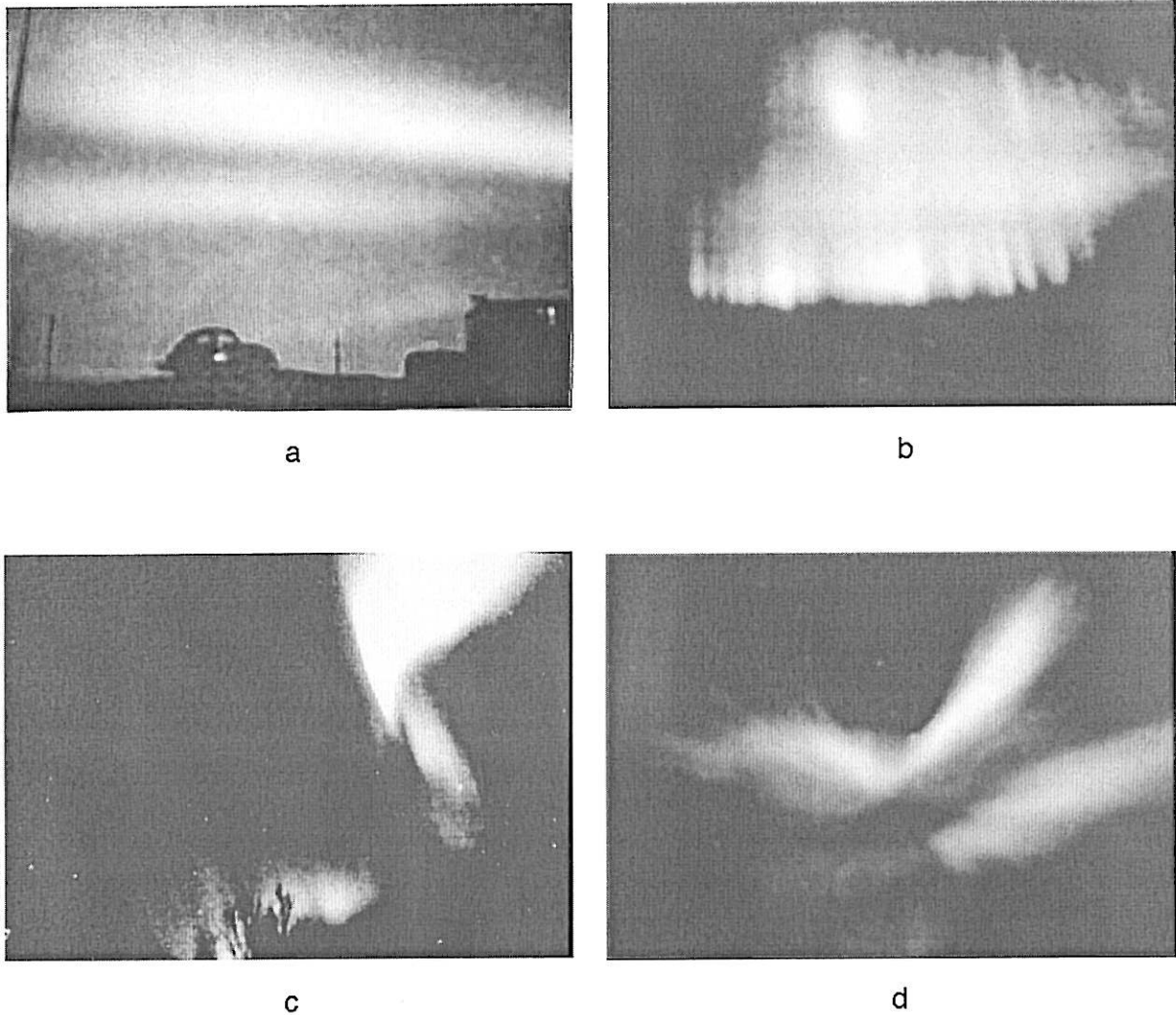


Figure 4.5.2. Examples of auroral forms (Eather 1980): a) Homogeneous arcs, b) Arcs with ray structure, c) Curtains, d) Corona.

Occurrence

Auroras occur mainly in two annular regions around the geomagnetic poles, the northern and southern *auroral zones*. Their locations are shown in Fig. 4.5.3. They are defined as the areas where in *average over a long time* the probability of occurrence of aurora is the greatest.

The geographic distribution *at a given instant in time* is, however, a different one and represented by the *auroral oval*. This is also an annular region but eccentric with respect to the magnetic pole, namely displaced toward the nightside as shown

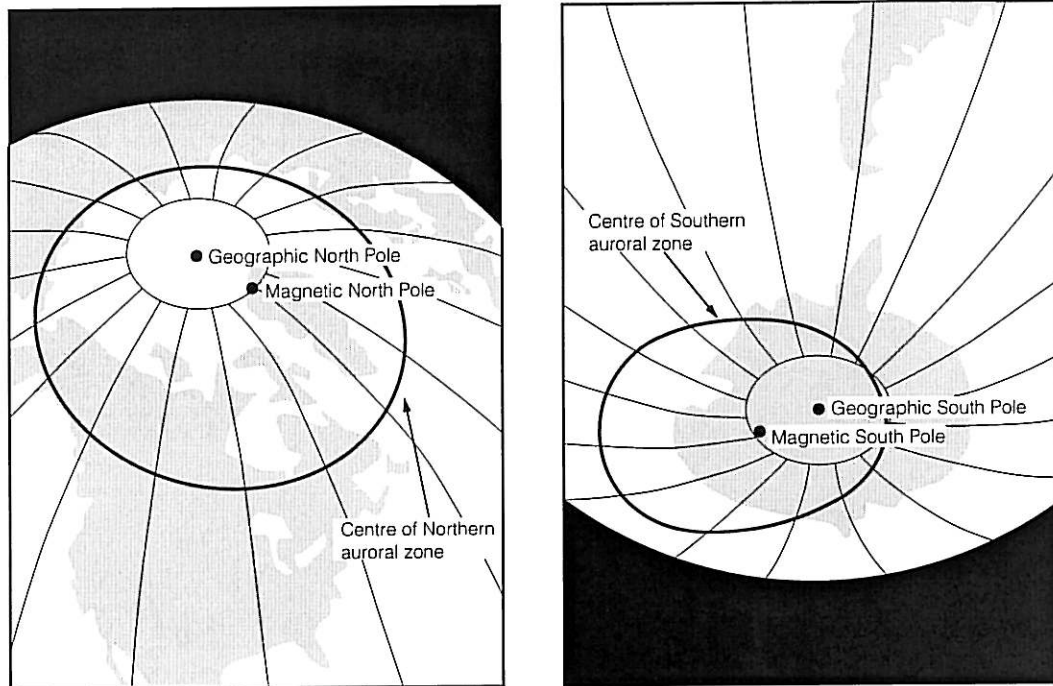


Figure 4.5.3. The northern and southern auroral zones (after Fälthammar 1972).

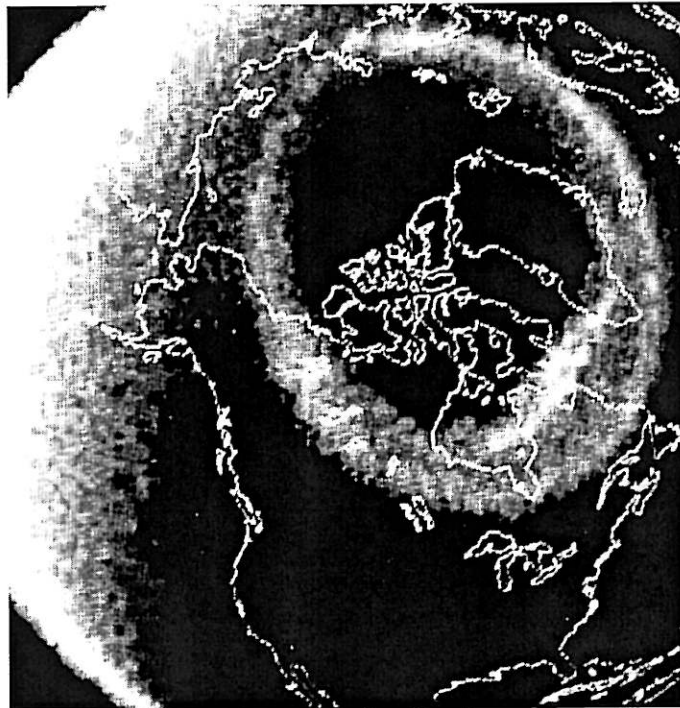


Figure 4.5.4. The northern auroral oval photographed (in ultraviolet light) from the American satellite Dynamics Explorer at a time when the USA is located under the nightside of the oval (Frank och Craven 1988. Fig. 3).

in Fig. 4.5.4. This eccentric oval is essentially fixed in space while the Earth rotates under it.

In the auroral oval, and between the auroral oval and space, electric currents of the order of magnitude 1 MA flow. The auroral current system is shown schematically in Fig. 4.5.6.



Figure 4.5.5. Aurora north of Scandinavia photographed with the American satellite DMSP (after Eather 1980).

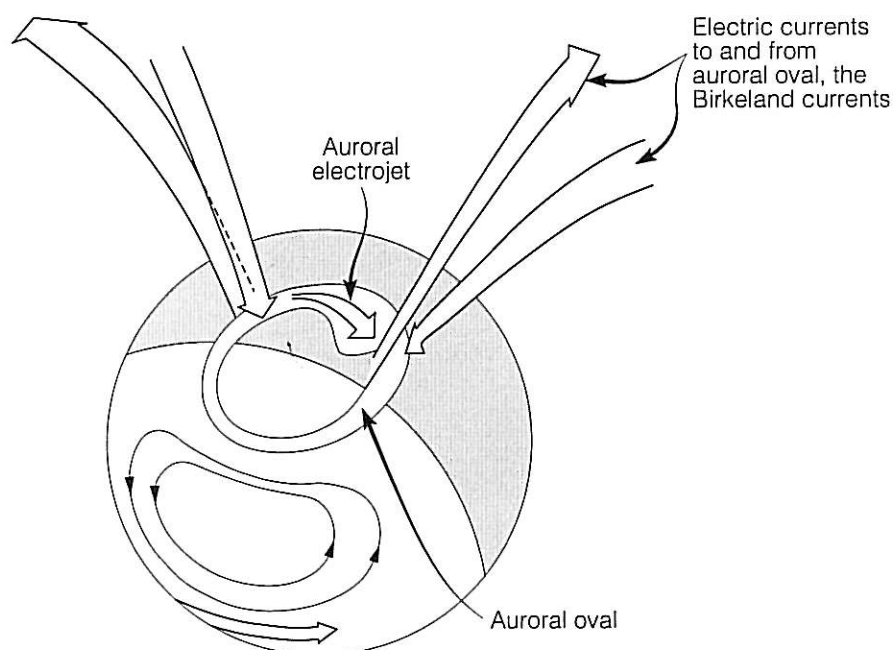


Figure 4.5.6. Sketch of the auroral electric current system (after Lanzerotti 1988). The currents that flows along the magnetic field lines to and from the auroral oval are called Birkeland currents

In contrast to the auroral zone, which is fixed in magnitude and shape (since it represents a long time average), the auroral oval varies both in radius and width. The higher the geomagnetic activity the wider and broader the auroral oval becomes.

During quiet conditions it shrinks, but never to zero. Within the auroral oval the individual auroral forms move, usually westward, in the early night and eastward in late night. These motions reflect a general flow of the magnetospheric plasma in the direction toward the Sun, associated with a general electric field directed from the morning side to the evening side of the magnetosphere.

During magnetospheric substorms (Section 4.4) the distribution of the aurora undergoes rapid changes in the ways illustrated in Fig. 4.5.7.

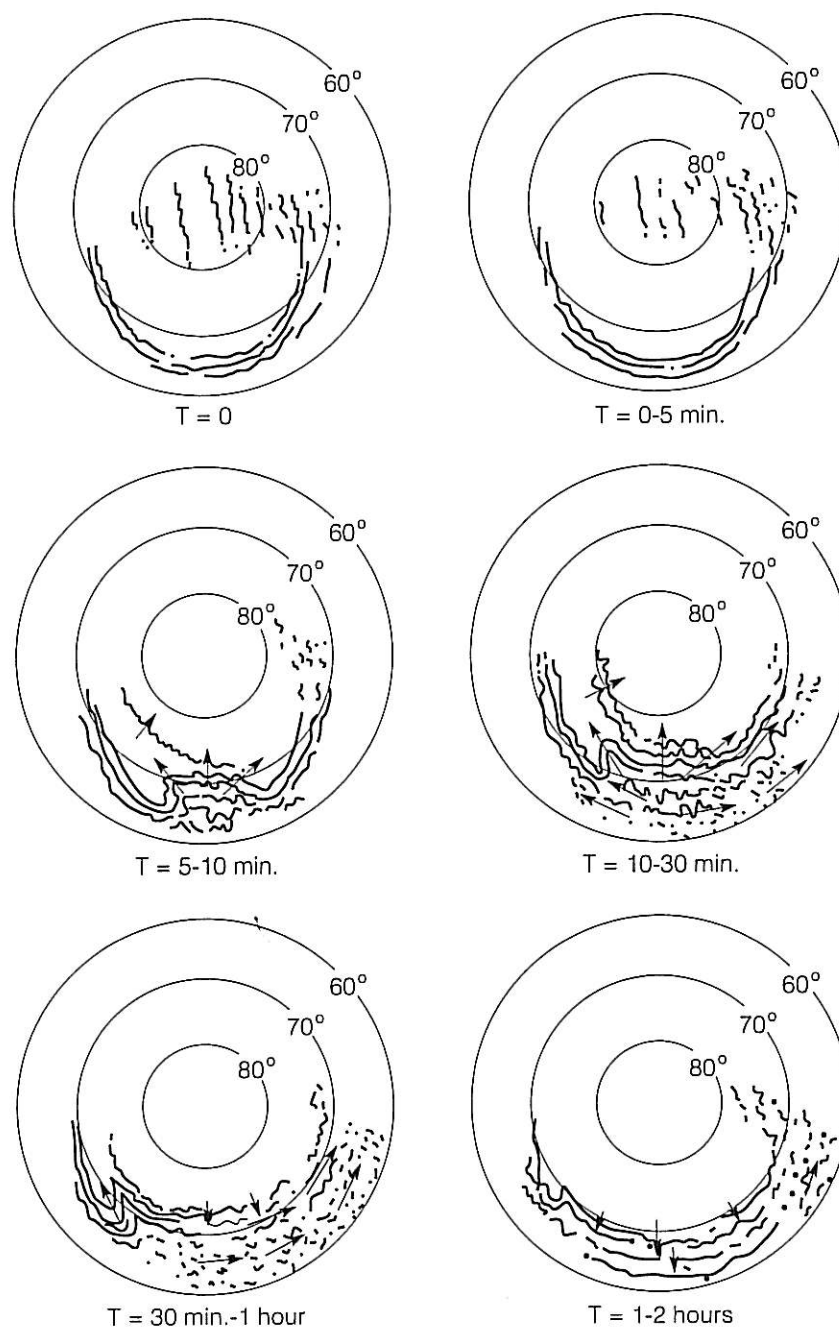


Figure 4.5.7. Schematic representation of the variations in the distribution of auroras during a substorm (Akasofu 1964).

A typical substorm can be described as follows. Before the substorm the aurora mainly forms quiet arcs. The dynamic phase starts with southernmost arc increasing in intensity and then after 5 to 10 minutes, “breaking-up”, *i.e.* intensifying, spreading and undergoing rapid motions, at the same time as the whole oval widens towards lower latitudes. Eventually the aurora gets fragmented, weakened and finally retreats to its previous distribution. Also the individual auroral forms undergo dramatic changes as shown in Fig. 4.5.8. minutes

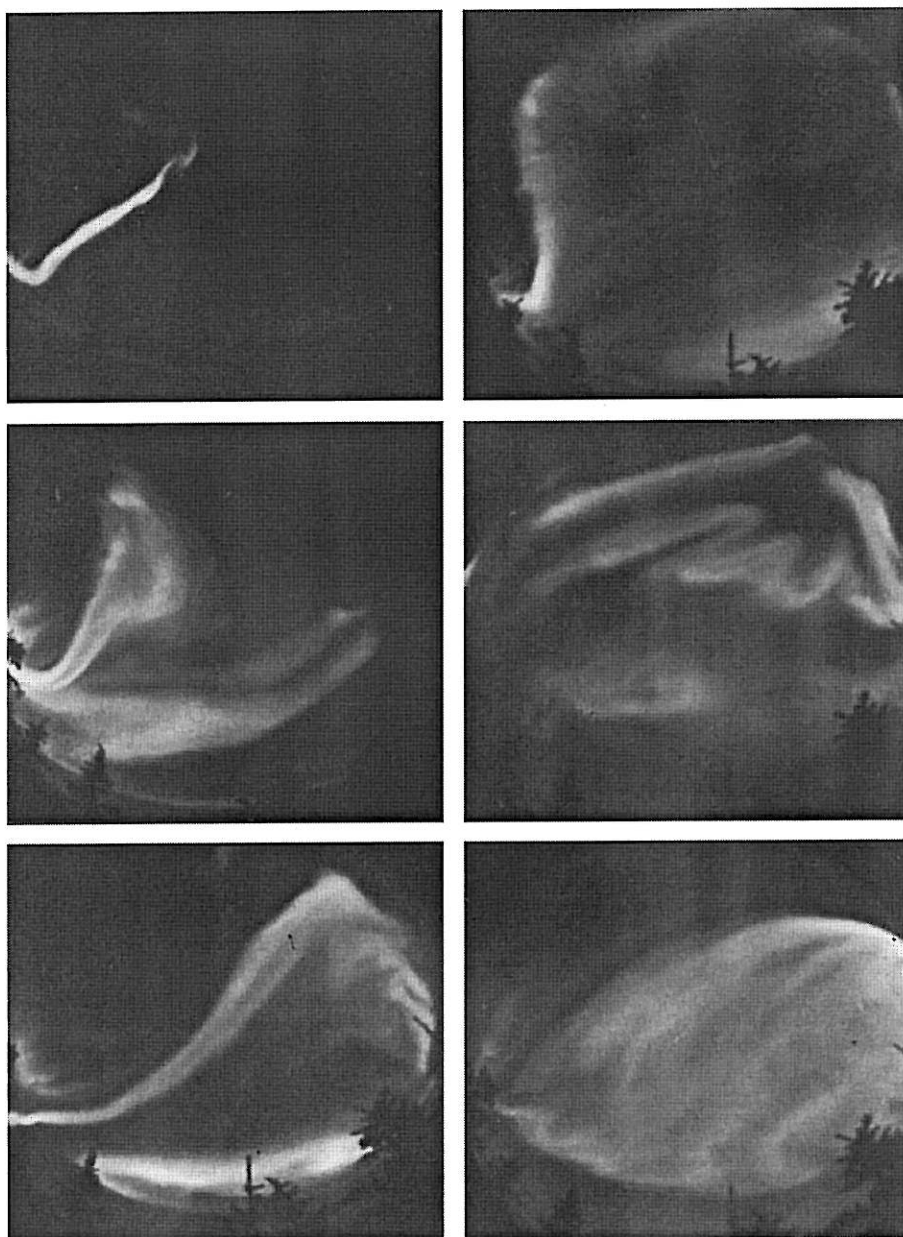


Figure 4.5.8. Aurora photographed from the ground with a fish-eye-lens. The time interval between two subsequent pictures is 4 minutes (Hones 1986).

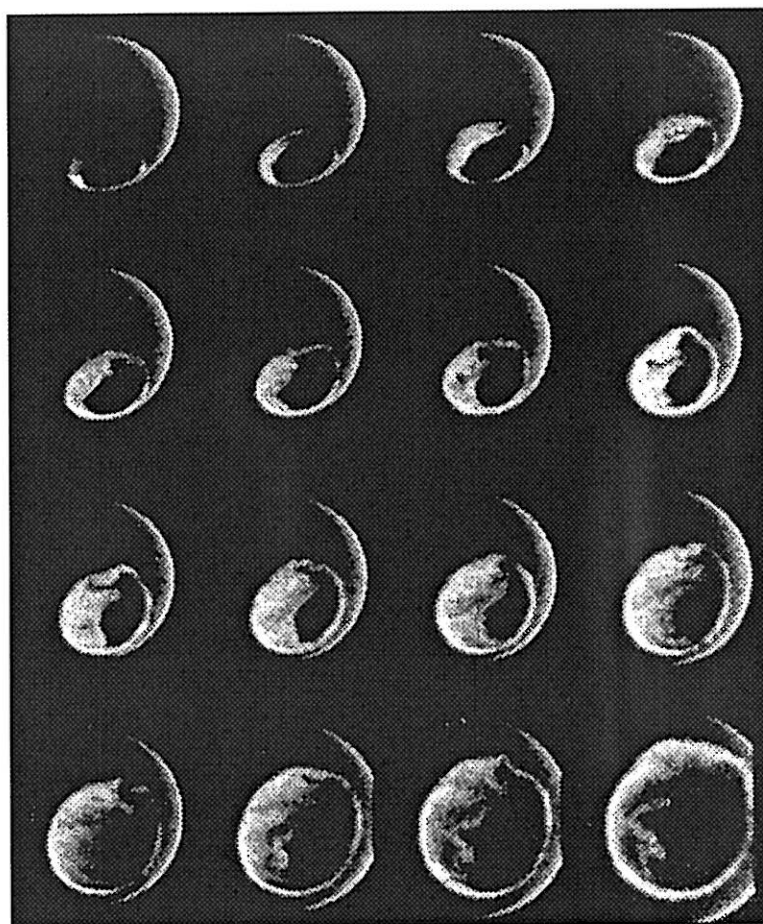


Figure 4.5.9. Sequence of UV-pictures from the satellite Dynamics Explorer, showing the variations of the auroral oval during a substorm (Frank and Craven 1988).

With the American satellite Dynamics Explorer it has been possible to directly photograph the auroral ovals in their entirety and their variation during a substorm. An example is shown in Fig. 4.5.9.

Physics of the Aurora

The direct cause of the aurora is that energetic particles, mainly electrons with energies of a few k eV, precipitate into the upper atmosphere and excite its molecules, which in turn emit light. This has been known for a long time.

The difficult problem concerning the aurora is, however, how these particles get their energy. Long ago Alfvén (1958) proposed that the auroral electrons get their energy by falling through an electric potential structure, electric double layer (cf. below) with a potential drop of a few kV. At that time it was, however, considered unthinkable that electric fields of this kind could occur in space. It was thought that plasma in space, which is electrically conducting, would immediately “short-circuit” such a potential drop. Alfvén’s suggestion was therefore not taken seriously.

When instruments carried on sounding rockets and satellites made it possible to conduct direct measurements of the particles that cause the aurora it was, however, found that their angular distribution and energy spectra strongly indicated that they had indeed undergone electrostatic acceleration (McIlwain 1960, Albert 1967, Evans 1968). Since then an increasingly overwhelming amount of data from rocket and satellite experiments have given a further support of the conclusion that Alfvén's idea was essentially correct, and that electrostatic acceleration along the geomagnetic field plays an essential role for the auroral process, even if the process in its entirety is complex and also includes various kinds of wave-particle interaction (Fälthammar 1983, 1985). The observations include both electrons accelerated downward toward the atmosphere and ions accelerated upwards and in addition "active" experiments with artificial ion clouds. The conclusion is also supported by direct measurements of electric fields, even if the component most important in this context, the one parallel to the magnetic field, still is technically very hard to measure.

Since charged particles move freely along magnetic field lines, it is the electric field component *along* the magnetic field that is particularly important for particle acceleration. It is, however, also this component that according to classical plasma theory must be zero ("the short-circuiting effect"). We know, however, now that there are mechanisms which make it possible for this component to be different from zero and sufficiently large to be of importance for an acceleration. For a review of possible mechanisms, see Fälthammar (1978) and for reviews of observations of electric field parallel to the magnetic field see Fälthammar (1985, 1986).

The initial result of electrostatic acceleration is a velocity distribution which is unstable and gives rise to various kinds of plasma waves, which modify the distribution.

The first identified mechanism for supporting electric fields along magnetic field lines, known as parallel electric fields, was a phenomenon called *anomalous resistivity*. It arises because the electric current drives plasma instabilities, which lead to intense fluctuating electric fields ("turbulent" wave fields) which strongly impede the motion of the current-carrying electrons. The result can be an effective resistivity many thousand times larger than what would be calculated from classical formulas based on ordinary binary collisions.

It has been found later that a very different phenomenon, called *electric double layer*, seems to be more important than anomalous resistivity. The electric double layer, too, is the result of plasma instabilities, but with an entirely different end state, namely that of one or more local potential drops with an extent of only some tenths of Debye lengths (cf. Section 1.3), which in the ionosphere can be as little as a few decimetres. The double layer phenomenon is very difficult to treat theoretically. A promising tool that has been used lately is numerical simulations using supercomputers. Experimentally the phenomenon is well-known and has also had technical importance (long ago in mercury rectifiers and more recently in certain laser discharges).

It is customary to distinguish between *weak* and *strong* double layers. The former has a potential drop which is comparable with the voltage equivalent of the electron temperature, *i.e.* kT_e/e , while the latter has a voltage drop which exceeds this value by a large factor.

Because of their small spatial dimensions and because they can vary strongly in time the electric double layers are hard to observe directly in space, but a number of indications of their existence have been found. One kind of weak double layers have, however, been directly observed. It has been possible to observe them because they are numerous and move rapidly along the magnetic field lines so that they have a great probability of passing and be observed by a satellite. They were first discovered by the American satellite S3-3 (Temerin *et al.* 1982) and have later been studied by the Swedish satellite Viking (Boström *et al.* 1987). Although each individual double layer has a small voltage drop (a fraction of a volt) they are so many that together they are estimated to be able to support a total voltage drop of several kV (Hudson *et al.* 1983).

Fig. 4.5.10. illustrates schematically how the acceleration of the auroral particles is now generally envisaged. The electric equipotentials, indicated as solid lines in the figure, are not stationary but exhibit rapid irregular variations in time and space. Associated with this basic process are a number of complicated and still largely unknown processes, which lead to the great complexity characterizing the auroral phenomenon.

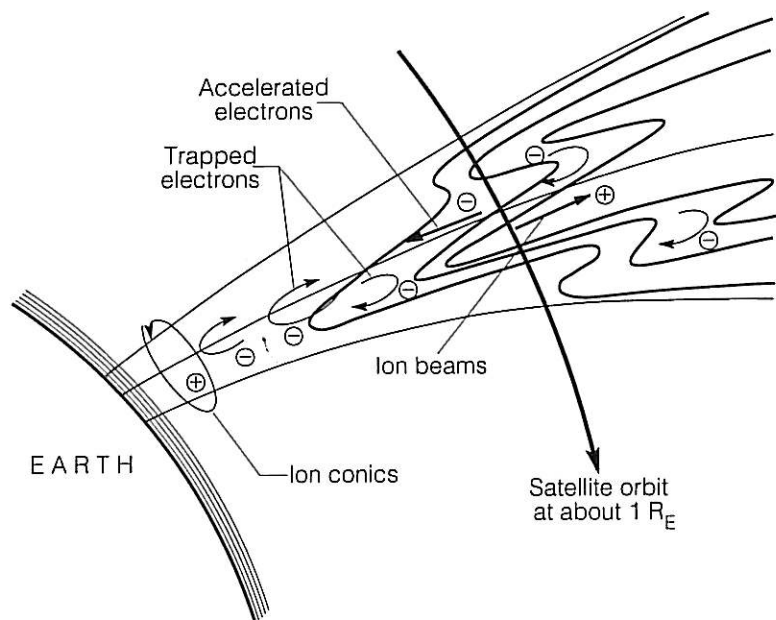


Figure 4.5.10. The acceleration of auroral particles is still incompletely understood. It is generally considered that electric fields with components along the magnetic field and varying irregularly in time and space, play an important role. (Chiu *et al.* 1983).

4.6. Mathematical Models of the Geomagnetic Field

The Geomagnetic Field at the Earth's Surface

The geomagnetic field at the Earth's surface is approximately a dipole field. The field lines from a magnetic dipole has the simple form previously illustrated in Fig. 1.2.1a.

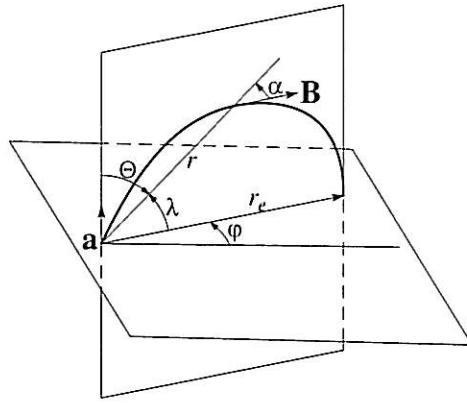


Figure 4.6.1. Geometry of the dipole field.

In spherical coordinates, r , θ , ϕ (see Fig. 4.6.1.) the components of a dipole field can be written

$$B_r = B_p \left(\frac{R_e}{r} \right)^3 \cos \theta \quad (4.6.1a)$$

$$B_\theta = \frac{B_p}{2} \left(\frac{R_e}{r} \right)^3 \sin \theta \quad (4.6.1b)$$

$$B_\phi = 0 \quad (4.6.1c)$$

and the magnitude of the field strength

$$B = |\mathbf{B}| = \sqrt{B_r^2 + B_\theta^2 + B_\phi^2} = B_p \left(\frac{R_e}{r} \right)^3 \frac{\sqrt{1 + 3 \cos^2 \theta}}{2} \quad (4.6.2)$$

where B_p is the polar magnetic field at the radius R_e .

The component equations (4.6.2) can be combined into the vector expression

$$\mathbf{B} = -\text{grad} \Phi \quad (4.6.3)$$

where the *scalar potential* Φ is given by the expression

$$\Phi = \left(\frac{\mu_0 \mathbf{a} \cdot \mathbf{r}}{4\pi r^3} \right) \quad (4.6.4)$$

the vector \mathbf{a} represents the *dipole moment* in terms of magnitude ($a = 2 \pi B_p R_e^3 / \mu_0$) and direction, and \mathbf{r} is radius vector from the dipole to the field point.

At the Earth's surface, *i.e.* at $r = R_e = 6378.4$ km, the magnetic field strength at the magnetic poles is $B_p = 62 \mu\text{T}$ (0.62 Gauss) and at the magnetic equator $B_p/2 = 31 \mu\text{T}$ (0.31 Gauss). Thus the magnetic dipole moment is $a = 8 \cdot 10^{22} \text{ Am}^2$. The magnetic pole in the Earth's northern hemisphere is a magnetic south pole, and vice versa.

For an accurate mathematical model of the geomagnetic field at the Earth's surface one can use a spherical harmonic series expansion, where the scalar magnetic potential Φ in the equation (4.6.3) is replaced by a scalar potential, Ψ , given by the expression

$$\Psi = R_e \sum_n \left(\frac{R_e}{r} \right)^{n+1} \cdot \sum_{m=0}^{m=n} (g_n^m \cos m\phi + h_n^m \sin m\phi) P_n^m(\cos \theta) \quad (4.6.5)$$

where g_n^m and h_n^m are empirically determined multipole coefficients and P_n^m are the associated Legendre functions.

The associated Legendre functions form a complete, orthogonal system of functions, and an expression of the form (4.6.5) can, if a sufficient number of terms are included, represent with arbitrary accuracy any magnetic field with only internal sources. The coefficients were originally determined from earthbound measurements. In more recent years satellites in low circular orbits have been used for determining them. The latest and most accurate mapping of the geomagnetic field near the Earth's surface has been made the satellite project Magsat.

Field Lines

The field lines of a magnetic field (sometimes inappropriately called lines of force) are lines which in every point are tangent to the local magnetic field vector. The field lines of an arbitrary magnetic field, \mathbf{B} , can therefore be calculated from the differential equations defined by

$$d\mathbf{s} = \text{konst. } \mathbf{B} \quad (4.6.6)$$

where $d\mathbf{s}$ is the vector line element of the field line. In Cartesian coordinates this equation can be written

$$dx/B_x = dy/B_y = dz/B_z \quad (4.6.6b)$$

and in spherical coordinates

$$dr/B_r = r d\theta/B_\theta = r \sin \theta d\phi/B_\phi \quad (4.6.6c)$$

Applied to the dipole field (4.6.1), (4.6.6c) gives

$$dr/B_r = r d\theta/B_\theta$$

or explicitly

$$dr/\cos \theta = r d\theta / \frac{1}{2} \sin \theta$$

i.e.

$$dr/r = 2\cos\theta d\theta / \sin\theta$$

which can be integrated to

$$\log r = 2 \log \sin \theta + \text{konst.}$$

and written

$$r = \text{konst.} \sin^2 \theta \quad (4.6.7)$$

which is the equation for the field lines of a magnetic dipole

The Magnetic Field in the Magnetosphere

With increasing distance from the Earth's surface the higher terms in (4.6.5) diminish, and the dipole term becomes dominant. At the distance of a few Earth radii, however, new and increasingly large deviations from the dipole field occur. These are due to external currents and can therefore not be described by the equation (4.6.5). The scalar potential must then have the more general form

$$\begin{aligned} \Psi = R_e \sum_n \left\{ \sum_{m=0}^{m=n} \left(\frac{R_e}{r} \right)^{n+1} (g_n^m \cos m\phi + h_n^m \sin m\phi) P_n^m(\cos \theta) \right. \\ \left. + \sum_{m=1}^{m=n} \left(\frac{r}{R_e} \right)^n (\gamma_n^m \cos m\phi + \kappa_n^m \sin m\phi) P_n^m(\cos \theta) \right\} \quad (4.6.8) \end{aligned}$$

where the first sum is identical with that in (4.6.5) (all terms of which decrease with increasing r). The new terms (all of which grow with r), represent a contribution from external currents.

However, in the outer magnetosphere not only internal and external sources come in to play but also currents in the magnetospheric plasma itself. Then a scalar magnetic potential does not apply.

A number of mathematical models have been developed on the basis of comprehensive satellite-borne magnetic field measurements in the magnetosphere. Since the real magnetic field varies with time, the parameters of the models have to be adjustable so as to apply in a given situation, e.g. for a given value of the magnetopause stand-off distance.

Among the most comprehensive magnetic field models can be mentioned the one developed by Tsyganenko (1987, 1989). (In these papers are also given references to other models.). Fig. 4.6.2 shows an example of the field lines of the magnetic field according to the Tsyganenko model.

Even if the newer theoretical models give a rather good representation of the local *magnetic field vector* at a given point under given conditions, *magnetic field lines* calculated from these models are still subject to considerable uncertainty. The reason is that when the field line equation for (4.6.6) is integrated, the deviations of the model from the real field can accumulate, and the resulting error in the path of the field line can be considerable. This certainly makes it difficult to relate phenomena observed in the ionosphere to their counterparts in the outer magnetosphere.

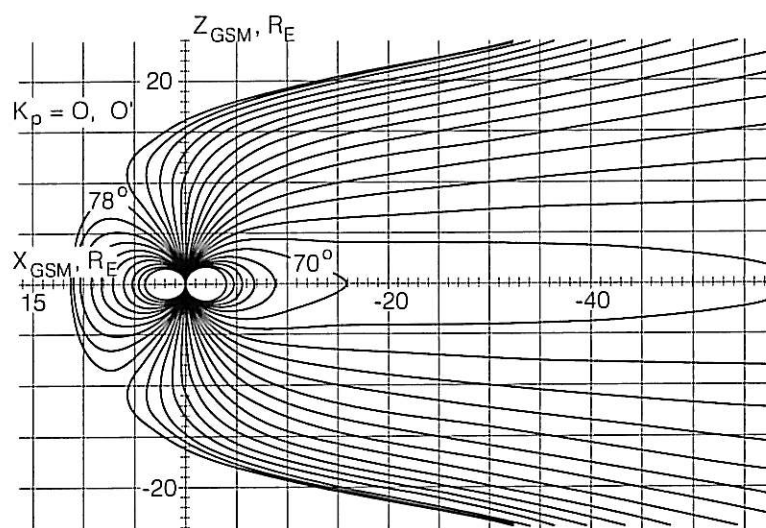


Figure 4.6.2. The shape of the magnetosphere according to Tsyganenko's (1987) model.

4.7. Field Transformations

According to the theory of relativity, both electric and magnetic fields depend on the state of motion of the observer. If an observer, at rest relative to a coordinate system S , observes an electric field \mathbf{E} and a magnetic field \mathbf{B} , an observer at rest relative to a different coordinate system, S' , will observe a different electric field \mathbf{E}' and a different magnetic field \mathbf{B}' . If the system S' moves relative to S with the velocity \mathbf{u} (which is assumed to be constant in order that the special theory of relativity should be valid), the following relations hold between the field observed components transverse to the direction of relative motion

$$\mathbf{E}' = \frac{\mathbf{E} + \mathbf{u} \times \mathbf{B}}{\sqrt{1 - u^2/c^2}} \quad (4.7.1)$$

$$\mathbf{B}' = \frac{\mathbf{B} - (\mathbf{u}/c^2) \times \mathbf{E}}{\sqrt{1 - u^2/c^2}} \quad (4.7.2)$$

(Panofsky and Phillips 1962, p. 330). In space physics it is usually true that the relative velocity, u , between different observers is such that $u \ll c$ and $E \ll cB$. Therefore the transformations formulas (4.7.1) and (4.7.2) are simplified to:

$$\mathbf{E}' = \mathbf{E} + \mathbf{u} \times \mathbf{B} \quad (4.7.3)$$

$$\mathbf{B}' = \mathbf{B} \quad (4.7.4)$$

i.e. the *electric*, but not the *magnetic* field depends on the state of motion of the observer. In other words, while the magnetic field is unique, it is meaningless to speak of the electric field if one does not also define the state of motion of the coordinate system to which it is referred. This is important in the magnetosphere where one sometimes wants to use a coordinate system following the rotation of the Earth and sometimes a non-rotating (inertial) system. Examples of this will be given later.

The non-relativistic formulas (4.7.3) and (4.7.4) can easily be derived in the following way.

Consider an electrically charged particle with mass m and charge q (including sign) which (relative to the coordinate system S) moves with a velocity \mathbf{v} in the presence of an electric field \mathbf{E} and a magnetic field \mathbf{B} . It is then acted on by an electric force $\mathbf{F}_E = q\mathbf{E}$ and a magnetic force $\mathbf{F}_m = q\mathbf{v} \times \mathbf{B}$, *i.e.* its equation of motion is

$$m \frac{d\mathbf{v}}{dt} = q\mathbf{E} + q\mathbf{v} \times \mathbf{B} \quad (4.7.5)$$

Let, as before, S' be another coordinate system moving relative to S with a velocity \mathbf{u} .

If \mathbf{v}' denotes the velocity vector of the particle in the coordinate system S' then

$$\mathbf{v} = \mathbf{u} + \mathbf{v}' \quad (4.7.6)$$

and since \mathbf{u} is constant,

$$\frac{d\mathbf{v}'}{dt} = \frac{d\mathbf{v}}{dt} \quad (4.7.7)$$

If (4.7.6) and (4.7.7) are substituted into (4.7.5) one finds

$$\begin{aligned} m \frac{d\mathbf{v}'}{dt} &= q\mathbf{E} + q(\mathbf{u} + \mathbf{v}') \times \mathbf{B} \\ &= q(\mathbf{E} + \mathbf{u} \times \mathbf{B}) + q\mathbf{v}' \times \mathbf{B} \end{aligned} \quad (4.7.8)$$

On the other hand, it is true by definition that in the system S'

$$m \frac{d\mathbf{v}'}{dt} = q\mathbf{E}' + q\mathbf{v}' \times \mathbf{B}' \quad (4.7.9)$$

Comparison between (4.7.8) and (4.7.9) immediately gives the transformation formulas (4.7.3) and (4.7.4).

Examples

A case where the field transformations have practical importance is in the context of measuring electric fields in space by means of rockets and satellites. A *sounding rocket* launched from a high latitude rocket range, for example ESRANGE, moves with a velocity whose component transverse to the magnetic field (about 40 μT) is of the order of a couple of hundred m/s. The electric field, \mathbf{E}' , measured by the rocket instrument, therefore differs from the electric field, \mathbf{E} , in a coordinate system fixed to the Earth by about $200 \cdot 0.4 \cdot 10^{-4} \text{ V/m} = 8 \text{ mV/m}$. As, in a quiet ionosphere, \mathbf{E} itself can be less than this, the difference between \mathbf{E}' and \mathbf{E} is very important. The importance of the field transformations becomes even more pronounced in the context of measurements by means of satellites, because of their greater velocity (7 km/s or more). At high latitudes, where the magnetic field forms a large angle with the satellite orbit, the term $\mathbf{u} \times \mathbf{B}$, *i.e.* the difference between the field measured by the satellite and the field measured in a system fixed to the Earth, can be over 200 mV/m. This means that in most cases the field measured by the satellite is mostly due to the motion of the satellite itself. The electric field, \mathbf{E} , in the system fixed to the Earth is thus obtained by subtraction of one large vector ($\mathbf{u} \times \mathbf{B}$) from another, nearly equally large vector, \mathbf{E}' , and this puts stringent requirements on the precision of the measurement of \mathbf{E}' in order to obtain an acceptable precision in the determination of \mathbf{E} .

Another example has to do with the rotation of the Earth. In spite of the fact that this is not a uniform motion, the transformation of the electric fields can be made by means of local transformations according to (4.7.3). If \mathbf{E} denotes the electric field in a coordinate system fixed in space and \mathbf{E}' the field in a system following the rotation of the Earth, the following relation holds

$$\mathbf{E}' = \mathbf{E} + (\boldsymbol{\omega} \times \mathbf{r}) \times \mathbf{B} \quad (4.7.10)$$

where $\boldsymbol{\omega}$ is the rotation vector of the Earth and \mathbf{r} the radius from the Earth's center. At auroral latitudes the term $(\boldsymbol{\omega} \times \mathbf{r}) \times \mathbf{B}$ is of the order of magnitude 7 mV/m. This means, among other things, that when describing the electric potential distribution in the auroral ionosphere it is necessary to define whether it is related to a rotating (Earth-fixed) or non-rotating system.

Considering the electrical potential over the whole ionosphere of the Earth, one finds that the integral from the pole to the equator of the term $(\boldsymbol{\omega} \times \mathbf{r}) \times \mathbf{B}$ is as large as about 90 kV. Thus even if there were no electric fields in a coordinate system fixed to the Earth, an observer fixed in space (non-rotating) would find that the poles have a (positive) potential of about 90 kV relative to the equator.

4.8. Motion of Magnetically Trapped Particles

In Section 3.7 it was described how electrically charged particles move in a magnetic field, and how a *drift motion* arises under the influence of electrical or other forces or a gradient of the magnetic field strength. The various types of drift motions transverse to the magnetic field have been summarized in Fig. 3.7.2..

Here we will describe the motion in an inhomogeneous magnetic field, especially with application to magnetically trapped particles. (For a more complete treatment, see *e.g.* Alfvén and Fälthammar 1963 or Fälthammar 1973a.)

Gradient Drift

In the case of a homogeneous magnetic field the magnetic force on the particle integrated over one gyration, is precisely zero. In an inhomogeneous magnetic field, on the other hand, the integral has a finite value. If the variation is slow, one can, as already mentioned in Section 3.7, show that the average force, $\bar{\mathbf{F}}$, is given in terms of magnitude and direction by the expression

$$\bar{\mathbf{F}} = -\mu \text{grad } B \quad (4.8.1)$$

where μ is the diamagnetic (gyro)moment derived in Section 3.7, which is given by (3.7.4).

If we first consider the motion transverse to the magnetic field, we find that the average force (4.8.1) is the same as the particle would experience in an electric field $\bar{\mathbf{F}}/q$. We can therefore apply the formula (3.7.10) for drift motion in an electric field ($\mathbf{u} = \mathbf{E} \times \mathbf{B}/B^2$) and find that the drift motion due to a magnetic field gradient becomes

$$\mathbf{u}_\perp = -\frac{\mu}{q}(\text{grad } B) \times \frac{\mathbf{B}}{B^2} \quad (4.8.2)$$

If the magnetic field lines are curved, the particle will, as a result of its motion along the curved field lines be subject to a centrifugal force. If the velocity component along the magnetic field is v_\parallel , the centrifugal force is $F_c = mv_\parallel^2/R$, where R is the radius of the curvature of the field line, and gives rise to a *curvature drift*, or *centrifugal drift* similar to that which an electric field of the strength F_c/q would give, *i.e.*

$$u_{\perp c} = \frac{mv_\parallel^2}{RqB} \quad (4.8.3)$$

One can show that in a current free magnetic field ($\text{rot } \mathbf{B} \equiv 0$), the gradient drift and the curvature drift can be combined and written

$$\mathbf{u}_\perp = -\left(1 + \frac{2v_\parallel^2}{v_\perp^2}\right) \frac{\mu}{q}(\text{grad } B) \times \frac{\mathbf{B}}{B^2} \quad (4.8.4)$$

(Alfvén and Fälthammar 1963, p. 34)

Magnetic mirroring

The comparison with electric field holds true also along the magnetic field lines. In that direction the particle therefore performs an accelerated motion just as it would have done in an electric field F/q , *i.e.*

$$\left(\frac{d\mathbf{u}}{dt}\right)_{\parallel} = -\frac{\mu}{m}(\text{grad } B)_{\parallel} \quad (4.8.5)$$

How the net force in the direction along the magnetic field arises can be qualitatively seen from the fact that the gradient in the direction of the magnetic field by necessity implies that the magnetic field lines come closer to each other in this direction, *i.e.* the magnetic field lines converge, see Fig. 4.8.1. Magnetic field lines on opposite sides of the gyro circle are therefore not entirely parallel and give the magnetic force a finite component transverse to the plane of the gyro circle. The force is so directed that it tends to expel the particle from the region with stronger magnetic field. fig.sp. 4.8.1 Illustration of the origin of net force in the case of a magnetic field gradient

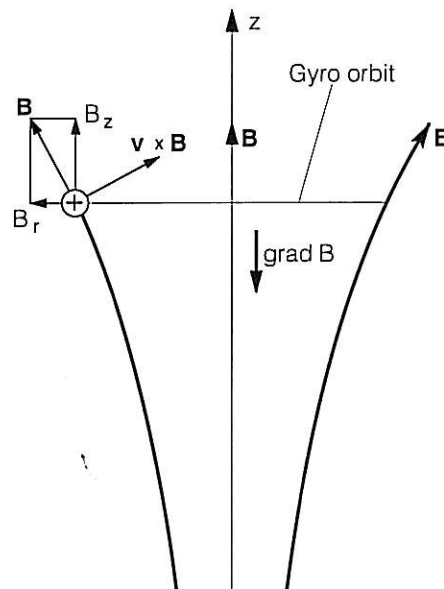


Figure 4.8.1. Illustration of the origin of net force in the case of a magnetic field gradient in the direction of the field.

If the magnetic field gradient is strong enough, the particle moving toward the stronger magnetic field will be decelerated and forced to turn. If the magnetic field lines converge in both directions from the position of the particle it can under certain circumstances be trapped between these regions of stronger magnetic field, which because of their ability of reversing the direction of motion of the particles are called

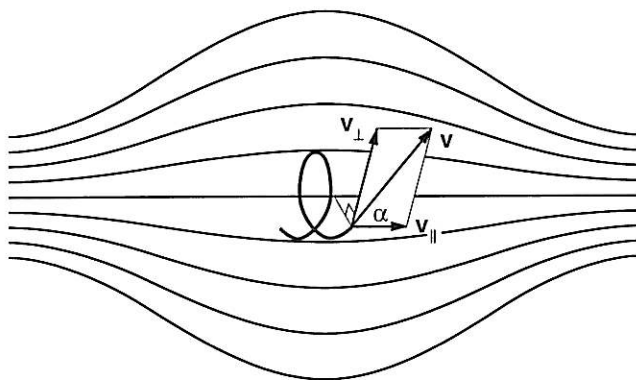


Figure 4.8.2. Particle trapped between magnetic mirrors.

“magnetic mirrors”, as illustrated in Fig. 4.8.2. The motion of a particle trapped between magnetic mirrors in the geomagnetic field is illustrated in Fig. 4.8.4.

The longitudinal oscillation

Particles trapped between two magnetic mirrors oscillate periodically back and forth between the magnetic mirrors (Fig. 4.8.3b) with a certain period, the *longitudinal period*, T_l .

The Azimuthal drift

A particle which is trapped longitudinally between two mirrors also usually performs an *azimuthal drift motion* consisting of a gradient drift and a curvature drift. If the drift orbit is closed, as illustrated in Fig. 4.8.3 (c), this motion, too, is periodic.

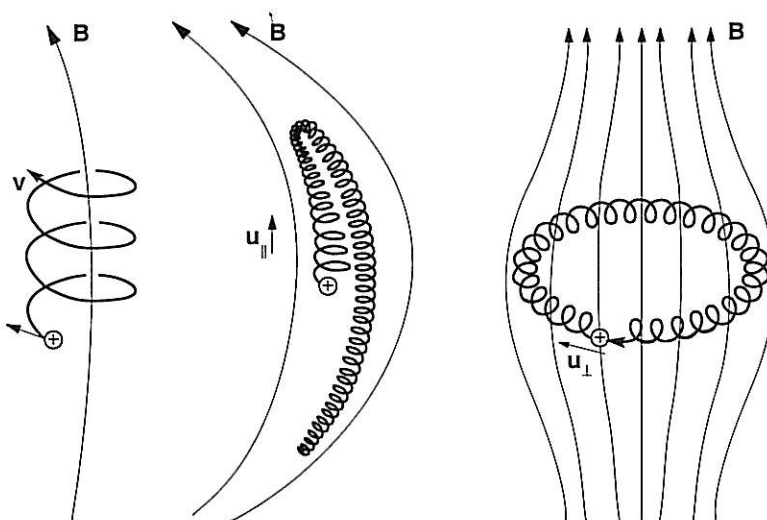


Figure 4.8.3. The three types of periodic motion gyration, longitudinal oscillation and azimuthal drift.

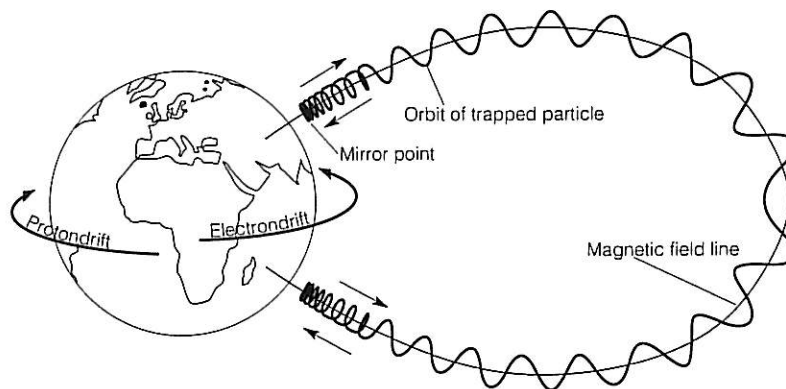


Figure 4.8.4. The orbit of a magnetically trapped particle in the geomagnetic field.

One can show that if l_c denotes a characteristic dimension of the spatial variation of the magnetic field, the three periods, the *gyroperiod*, T_g , the *longitudinal period*, T_l the *azimuthal period*, T_a , have the following relative magnitudes:

$$T_a : T_l : T_g = (l_c/\varrho)^2 : (l_c/\varrho) : 1 \quad (4.8.6)$$

where ϱ is the gyroradius.

Adiabatic invariance

An important concept in the theory of magnetically trapped particles is *adiabatic invariance*.

The classic example of an adiabatic invariant is the ratio of energy to frequency of an oscillating pendulum. If the length of the pendulum is changed slowly (so that the relative change during each individual oscillation period is small), this ratio is almost unchanged – even after the accumulated change of the pendulum length has become large – and the change in energy can be made arbitrarily small by changing the pendulum length sufficiently slowly. In this case the ratio of energy to frequency of the pendulum is an *adiabatic invariant*.

The Magnetic Moment – the First Adiabatic Invariant

It can be shown that the magnetic moment μ , which according to (3.7.4) is given by the expression

$$\mu = mv_{\perp}^2/2B \quad (4.8.7)$$

is an adiabatic invariant, also called the *first adiabatic invariant*.

No matter how much B changes, also v_{\perp} , and the corresponding kinetic energy, W_{\perp} change just in such a way that μ remains very nearly unchanged, *provided* that the change takes place so slowly that the relative change during each individual gyration period is small. This is true whether the change is caused by the magnetic field, B , varying in time or the particle drifting into a region with a different value of B . How strong this invariance is can be illustrated by the following theorem (see *e.g.* Northrop 1963) : when the “parameter of slowness”, δ (for example the

relative change of B per gyro period), tends to zero, the change in μ tends to zero *more rapidly than any finite power of δ* .

The longitudinal Invariant – the Second Adiabatic Invariant

Just as there is an adiabatic invariant corresponding to the periodic gyro motion, the magnetic moment, there exists an adiabatic invariant corresponding to the longitudinal oscillation, namely the *longitudinal invariant* or the *second adiabatic invariant*, J , which is defined by

$$J = \oint m v_{\parallel} ds \quad (4.8.8)$$

The condition for its conservation is that the magnetic mirror field changes very little during each individual longitudinal oscillation period.

The Flux Invariant – the Third Adiabatic Invariant

It can be shown that there also exists an adiabatic invariant corresponding to the periodic azimuthal drift, the *third invariant*, which is equal to the *magnetic flux*, Φ_a , enclosed by the drift orbit. It is therefore called the *flux invariant*. Like in the previous cases, the condition for invariance is that the changes are slow relative to the drift period.

Since, as is apparent from (4.8.6), the periods T_g , T_l and T_a are very different in length, the three adiabatic invariants are very differently sensitive to disturbances (Φ_a most, μ least). For magnetically trapped particles in the magnetosphere, it is often true that the first and second invariants, but not the third, are conserved.

Loss Cone

In order to determine quantitatively the condition for magnetic trapping we can use the fact that the magnetic moment, μ , is an adiabatic invariant. This means that as long as the magnetic field along the orbit of the particle changes slowly in the sense that the change is small during any individual gyro period, the magnetic moment retains its value even if the total change of the magnetic field is large. This is true whether the change of the magnetic field is due to a pure time variation or due to the particle moving in an inhomogeneous magnetic field.

If the angle between the velocity vector and the magnetic field, called *pitch angle*, is denoted by α and the magnitude of the velocity vector by v ($v = \sqrt{v_{\parallel}^2 + v_{\perp}^2}$), the expression for the magnetic moment (3.7.4) can be written

$$\mu = \frac{m v_{\perp}^2}{2B} = \frac{m v^2 \sin^2 \alpha}{2B} \quad (4.8.9)$$

Since the energy of the particle is conserved (because the force is transverse to the motion), the magnitude of the velocity, v , is constant too.

For a particle with the moment μ , which at a point with the magnetic field strength B has the pitch angle α , α will increase with the strength of the magnetic field. The strongest magnetic field, B_{max} , that the particle can penetrate into, is given by the condition that α reaches the value of 90 degrees (which means that the particle turns around), *i.e.*

$$\mu = \frac{mv^2 \sin^2 \alpha}{2B} = \frac{mv^2 \cdot 1}{2B_{max}}$$

$$B_{max} = B / \sin^2 \alpha \quad (4.8.10)$$

Conversely this means that if the magnetic mirror field has the maximum field strength, B_{max} , it can trap all particles which at the field strength B have a pitch angle larger than

$$\alpha = \alpha_{fl} = \arcsin \sqrt{B/B_{max}} \quad (4.8.11)$$

Particles with smaller pitch angle than α_{fl} will penetrate the mirror. The expression (4.8.11) thus defines a cone in velocity space, the *loss cone*, which separates trapped from non-trapped particles.

Adiabatic Acceleration

Since the expressions for the first two invariants contain the velocity of the particle, one can use the invariance to change the velocity of the particle and hence its energy. From (4.8.7) it follows that an increase of B leads to an increase of the velocity component v_{\perp} and thereby the corresponding part of the energy, W_{\perp} , since the adiabatic invariant μ is unchanged. Thus one can accelerate the gyrating particle simply by increasing B slowly. This is called *betatron acceleration* and is a form of *adiabatic acceleration*.

Similarly it is seen from (4.8.8) that if the magnetic field is changed in such a way that the distance between the magnetic mirrors, *i.e.* the length of the integration path in (4.8.8), decreases, v_{\parallel} must instead increase for the adiabatic invariant J to be unchanged. Thus one can accelerate a particle by deforming the magnetic field in such a way that the magnetic mirrors approach each other (or by using an electric field to force the particle to drift into the region where the magnetic mirrors are closer to each other). This is called the *Fermi acceleration* after the Italian nuclear physicist Fermi, who suggested this mechanism in order to explain how the cosmic radiation obtains its energy.

A third kind of adiabatic acceleration, corresponding to the third invariant, does not exist, since the expression for this invariant does not involve the energy of the particle.

	Period	Invariant	Acceleration
Gyration	T_g	μ	Betatron acceleration
Longitudinal oscillation	T_l	J	Fermi acceleration
Azimuthal drift	T_a	ϕ_a	---

Table 4.8.1. Periodicities, adiabatic invariants and adiabatic accelerations.

The periodic motions and the corresponding adiabatic invariants and adiabatic accelerations are summarized in Table 4.8.1.

When one first tried to describe the spatial distribution of the magnetically trapped particles in the Earth's radiation belts, it turned out to be extremely complicated, due to the deviation of the geomagnetic field from an ideal dipole shape. This problem was solved in a brilliant way by Carl McIlwain (1961, 1966), who introduced the coordinate system based on the adiabatic invariants. In this coordinate system the coordinate surfaces used are (1) the surfaces $B = \text{const}$ and (2) the surfaces that a particle generates through its longitudinal oscillation and azimuthal drift. The latter are labeled by a parameter, which in a dipole field is equal to the equatorial radius of the drift surface, measured in Earth radii. To each such L -value, as it is called, there corresponds an *invariant latitude*, Λ , given by

$$1 = L \cos^2 \Lambda \quad (4.8.12)$$

In a dipole field Λ is equal to the latitude where the field lines of the drift shell cut the Earth's surface.

This coordinate system is now routinely used everywhere as computer programs are available for transformation between the ordinary spatial coordinates and the B - L -system (cf. Fig. 4.8.5).

Radial Diffusion

The fact that fluctuations of the geomagnetic field, and – even more important – time-dependent electric fields, disturb the motion of magnetically trapped particles, causes a form of diffusion, which is a fundamental prerequisite for the existence of large parts of the radiation belts. The perturbations are usually sufficient to violate the third invariant but not the first two. The perturbed drift orbits do not entirely close but are displaced randomly outwards or inwards. The result is a

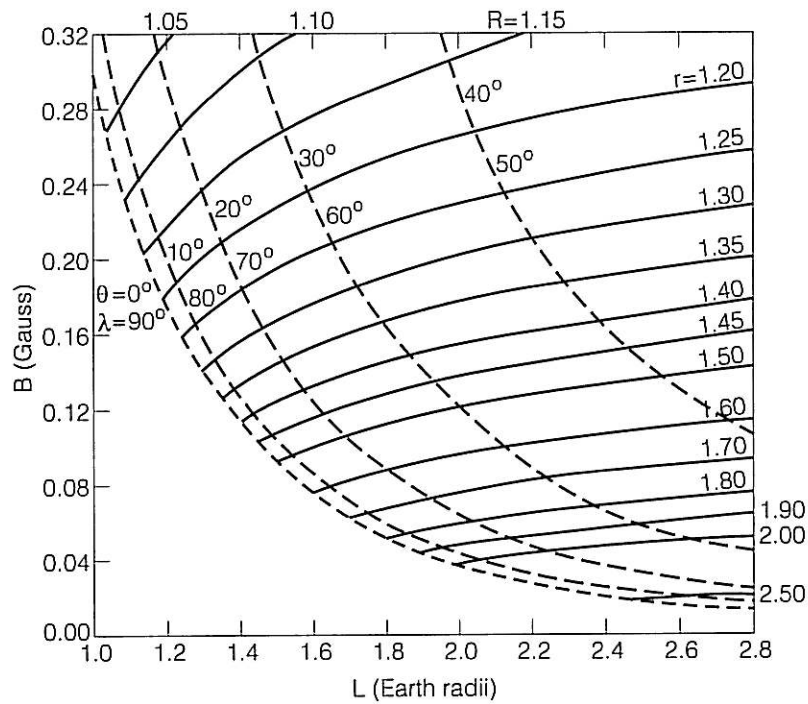


Figure 4.8.5. Relation between the B-L-coordinates and the spherical coordinates, r , and θ (McIlwain 1961).

“random walk” in the L -parameter, and hence a *radial diffusion*. Particles are supplied from the outside, while particles sufficiently close to the Earth are lost through absorption or other loss processes. Thus there is a source on the outside and a sink in the interior. This leads to a net inward transport of energetic particles, which in this way refill the radiation belts.

4.9. Other Magnetospheres

Like the Earth other magnetized planets and satellites are surrounded by magnetospheres. Magnetospheres also occur around other astrophysical objects. A brief survey will be given below.

Direct measurements by means of space probes have now been made near most of the planets in our own solar system. Among the most important of these space probes are the two Voyager probes. Both were launched from Cape Canaveral in Florida in 1977, Voyager 2 on the 20th of August and Voyager 1 on the 5th of September. (Voyager 1, which followed a faster trajectory, was the one to arrive first.) The Voyager probes, whose appearance and equipment are shown in Fig. 4.9.1, have functioned well for more than a decade, an impressive achievement.

Voyager 1 has, after passing Saturn, left the solar system and is on its way into interstellar space. Voyager 2 has made close encounters with Jupiter, Saturn and Uranus, and passed by Neptune on the 25th of August 1989, at 05.00 UTC. The orbits of the Voyager probes and their passage times are shown in Fig. 4.9.2.

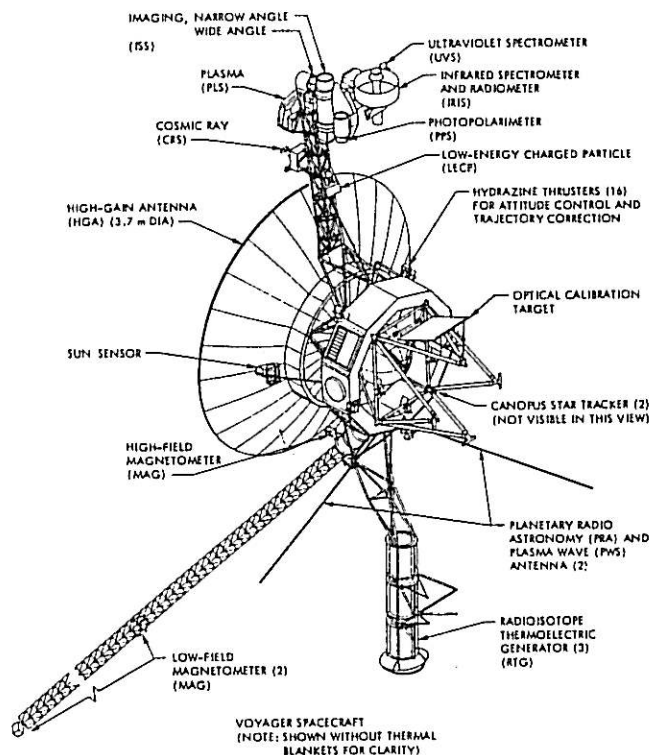


Figure 4.9.1. Structure and equipment of the Voyager space probes (Stone 1983).

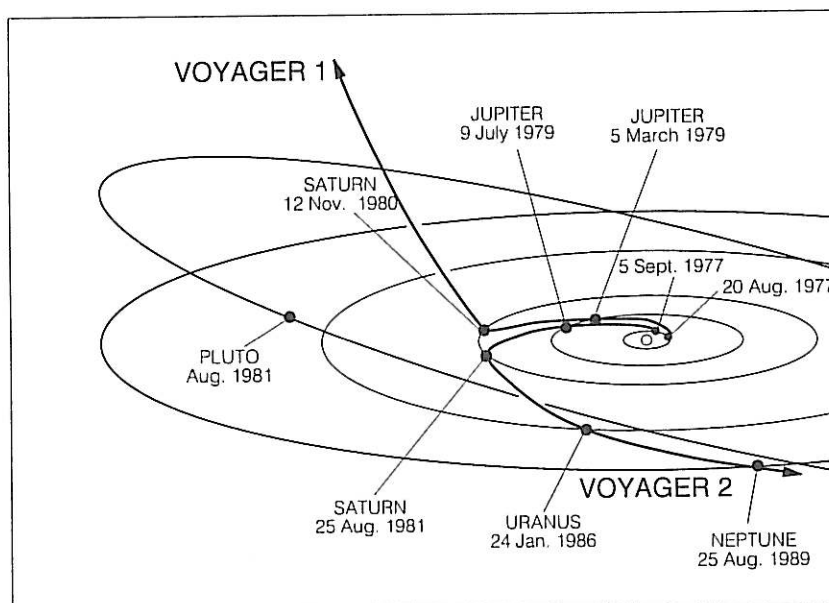


Figure 4.9.2. The orbits of the Voyager probes.

	Radius Earth radii	Spin period (days)	Equatorial field strength (μT)	Magnetic axis direction relative to spin axis	Polarity relative to Earth's	Typical magneto- pause distance (planetary radii)
Mercury	0.38	58.6	0.35	10^0	Same	1.1
Venus	0.95	243	< 0.03	-	-	1.1
Earth	1.0	1	31	11.5^0	Same	10
Mars	0.53	1.02	0.065	-	Opposite	?
Jupiter	11.18	0.41	410	10^0	Opposite	60-100
Saturn	9.42	0.44	40	$< 1^0$	Opposite	20-25
Uranus	3.84	0.72	23	60^0	Opposite	18-25
Neptune	3.93	0.74	20-150 ^{*)}	47^0	Opposite	26 ^{**)}

^{*)} The magnetic field differs greatly from a dipole field. The numbers represent maximum and minimum strength at the planetary surface

^{**) Based on single passage}

Table 4.9.1. Magnetic parameters of the planets (Lanzerotti and Krimigis 1985, Ness *et al.* 1989).

The main characteristics of the planetary magnetic fields are summarized in Table 4.9.1.

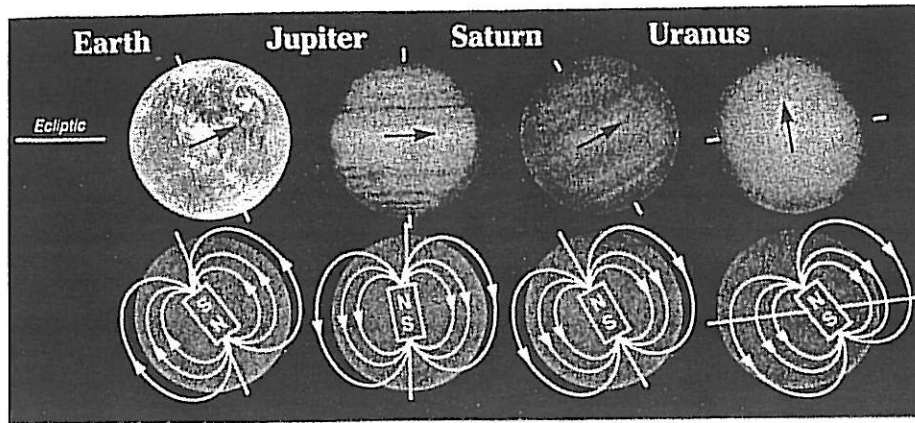


Figure 4.9.3. Orientations of the magnetic dipoles of the major planets (Lanzerotti and Uberoi 1989).

Fig. 4.9.3 shows the orientation of the magnetic dipoles of the Earth and of the planets Jupiter, Saturn and Uranus.

Mercury

The fact that Mercury, in spite of its slow rotation (rotation period equivalent to 58.6 Earth rotation periods) has a magnetic field which was discovered with the *Mariner* space probe that went to Venus and Mercury (Ness 1979). Because it is also located closer to the Sun, and therefore in a denser solar wind than the Earth, its magnetopause is located less than a planetary radius above its surface. Since Mercury does not have any appreciable atmosphere, it also lacks an ionosphere. At the passage through the magnetotail of Mercury, Mariner measured showers of electrons, both low energy electrons and those of energy in range 70 – 300 keV.

Venus

Venus has no intrinsic magnetic field, and therefore no magnetosphere of the kind that the Earth has. On the other hand it has an *ionosphere*, and the border between the solar wind and the ionosphere, the *ionopause*, is what is entered in the column “magnetopause” in Table 4.9.1. Just as there is a bow shock in front of a magnetopause there is also a bow shock in front of Venus’ ionopause. Venus also has a magnetotail (formed by magnetic field from the solar wind), which extends to at least 10 Venus radii and has a filamentary structure. The interaction of the solar wind with Venus is therefore similar to its interaction with comets and with the Saturn-satellite Titan, which has an atmosphere but no magnetic field.

Mars

There are still different opinions about the intrinsic magnetic field of Mars and about its possible magnetosphere. There are some indications that the solar wind interacts with an ionosphere in the upper parts of the thin planetary atmosphere. The absence of a strong magnetic field has been proposed as an explanation for the Mars atmosphere being so thin, the original atmosphere has largely been carried away by the solar wind.

Jupiter

The largest planet of the solar system has also the largest magnetosphere of the solar system. It extends 60 – 100 Jupiter radii in the direction toward the Sun and at least 5 astronomical units (about 11 000 Jupiter radii) on the night side. If Jupiter's magnetosphere were visible from the Earth, it would subtend a solid angle about ten times greater than the Sun, in spite of the fact that it is located more than four times as far away. (Even with its high velocity of about 10 km/s the space probe Voyager 2, needed 32 days to travel through parts of Jupiter's magnetosphere.)

The size of Jupiter's magnetosphere is not entirely a result of the strong planetary magnetic field. The magnetosphere contains large amounts of plasma, which in large parts of the magnetosphere follows the rotation of the planet. The pressure of the plasma and the centrifugal force from its rotation plays the most important role in the pressure balance with the solar wind.

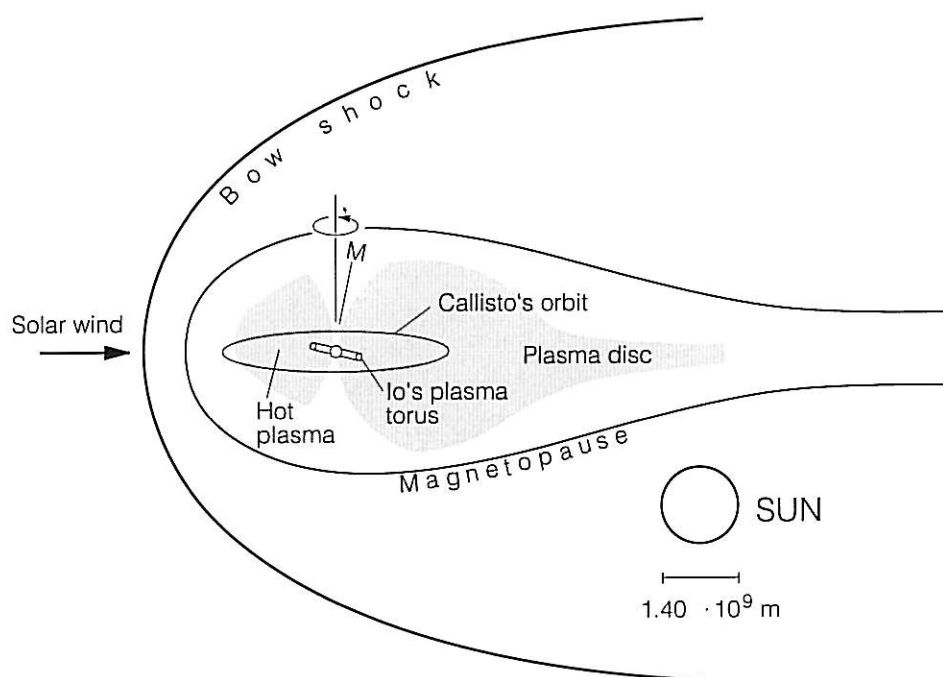


Figure 4.9.4. Sketch of Jupiter's magnetosphere (Lanzerotti and Uberoi 1989).

In the case of Jupiter, five of the satellites, among them the Galileian satellites (Io, Europa, Ganymede and Callisto) orbit within the magnetosphere. The magnetospheric plasma corotating with the planet sweeps past these satellites in the same way as the solar wind sweeps around the planets of the solar system. In this context the plasma phenomenon the “critical velocity” (described in Section 6.4) is considered to play a role for the formation of an ionosphere around the satellite Io (Cloutier *et al.* 1978).

The relative motion between the magnetospheric plasma and the satellite Io corresponds to an electric field with a potential difference of about 400 kV across the satellite. This field drives electric currents along the magnetic field lines between Io and Jupiter’s ionosphere. According to magnetic field measurements made with the space probe Voyager 1 these currents have a strength of about 1 MA.



Figure 4.9.5. Aurora (upper left corner) on the planet Jupiter photographed with the space probe Voyager 1 (Soderblom *et al.* 1979). The spots to the right are lightning.

The violent volcanic activity on Io injects on the average 1000 kg of gases per second and creates a ring of plasma near the orbit of Io (Io’s “*plasma torus*”). The magnetospheric density is in some regions so large (about 2 000 ion pairs per cm^{-3}) that the corotation is locally strongly reduced because of the load. Further out, where density again is lower, corotation is essentially maintained all the way to the magnetopause.

Even high-energy ions (30 keV to over 100 MeV) and electrons (a few keV to over 40 MeV) occur copiously in Jupiter’s magnetosphere and even leak out from it. For example, electrons leaking from Jupiter’s magnetosphere, constitute most of what until 1974 was considered to be the electron component below 30 MeV of the cosmic radiation (cf. chapter 9).

Intense fluxes of electrons above 10 MeV and ions above 100 MeV in the inner magnetosphere of Jupiter drive plasma instabilities, which give rise to intense *radio wave emission*, mostly in the 10 MHz-range, *decameter waves* (see Fig. 4.9.6) and make Jupiter a very strong *radio source*.

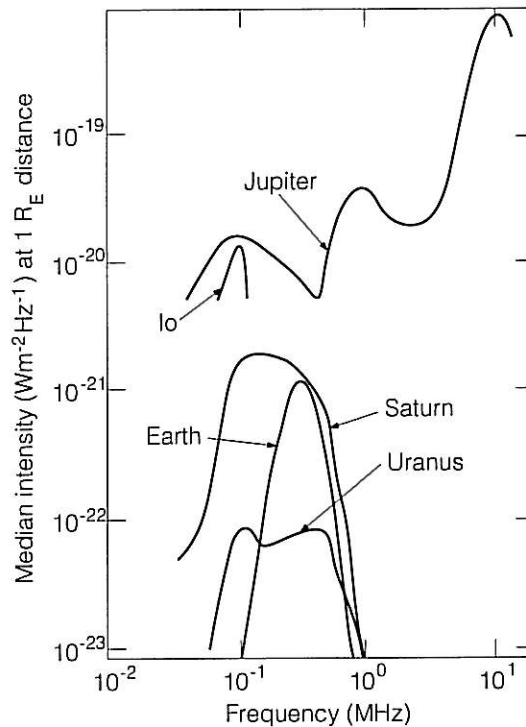


Figure 4.9.6. Radio emission from the planets (Lanzerotti and Krimigis 1985).

In essence the mass and energy balance in the magnetosphere of Jupiter is as follows: Io contributes $2 \cdot 10^{28}$ ion pairs/s (sulphur, oxygen, sodium, calcium) and a power of $2 \cdot 10^{12}$ W. Jupiter's upper atmosphere contributes 10^{28} ion pairs/s (hydrogen) and $3 \cdot 10^{14}$ W. The contribution of the solar wind is estimated at 10^{28} ionpairs/s (mostly hydrogen) and perhaps $3 \cdot 10^{13}$ W. Thus the dominating contributions to the energy balance come from the rotation of the planet (which, however, is not appreciably influenced by this energy loss). This implies a similarity between the magnetosphere of Jupiter and those of pulsars (Chapter 7). Part of the hot plasma of the magnetosphere is lost by way of a plasma wind which streams out along Jupiter's magnetotail. With a speed exceeding that of the solar wind this "planetary wind" transports about $2 \cdot 10^{27}$ ions/s and a power of about $2 \cdot 10^{13}$ W.

Also the aurora on Jupiter, which is believed to dissipate $2 \cdot 10^{13}$ W in the form of precipitating ions, is driven by the rotation of the planet.

Since plasma from the planet or its satellites is present practically all the way out to the magnetopause, one may say that the plasmasphere of Jupiter occupies essentially the whole magnetosphere.

Saturn

The magnetic field of Saturn is surprisingly weak, with a surface field strength about as great as that of the Earth, see Table 4.9.1. Voyager 1 and Voyager 2 have found plasma with energies of 20–55 eV in quantities which indicate that the plasma pressure is greater than the magnetic field pressure. If so, this means that the magnetosphere of the Saturn is more similar to Jupiter's than to the Earth's. (The relative plasma pressure is usually expressed in terms of the β -value: $\beta = p/(B^2/2\mu_0)$. Thus in the magnetosphere of Saturn β is greater than 1).

Also high energy magnetically trapped radiation (50 MeV protons) has been observed and is considered to originate from the neutron albedo of the Saturn atmosphere (cf. Chapter 9).

The physical processes in the magnetosphere of Saturn are further complicated by the fact that not only some of the satellites but also the rings of Saturn are located within the magnetosphere and give rise to both sources and sinks of particles and to interactions between plasma gas, dust particles and larger solid bodies. For example the A and B rings act as absorbers and stop energetic particles from diffusing closer to the planet (about in the same way as a "limiter" in a fusion plasma experiment).

In terms of chemical composition the magnetosphere of Saturn consists of mainly *hydrogen*, *oxygen* and *nitrogen*. (The oxygen derives mainly from sputtering at the icy surfaces of the satellite.) The nitrogen-dominated atmosphere of the Saturn satellite Titan is considered also to contribute to the plasma content of the Saturn magnetosphere. (On the dayside of Saturn, Titan is sometimes within, sometimes outside the magnetopause, depending on the strength of the solar wind, on the nightside it is always within.)

The aurora at Saturn, discovered through UV-measurements from Voyager 1, appears to have an intensity corresponding to that of the earthly auroras.

Uranus

Uranus deviates from all the other planets by its rotation axis being almost in the orbital plane. It had been expected that also the magnetic axis would be likewise oriented. But one of the discoveries made by Voyager 2 was that the dipole moment of Uranus forms an angle of almost 60 degrees with the planet's spin axis.

The magnetosphere of Uranus was found to be filled with hydrogen ions (and equally many electrons). The plasma density is not particularly high and the β -value low. There are indications of radiation belts close to the planet. But since Voyager 2 did not come sufficiently close to reach them, the indications are still unconfirmed. Voyager 2 detected UV-radiation that indicated auroral activity which (at the time) was somewhat weaker than on Saturn and on the Earth.

Neptune

Before the Neptune passage of Voyager 2, which took place on the 25th of August 1989, almost nothing was known about the magnetosphere of Neptune. A

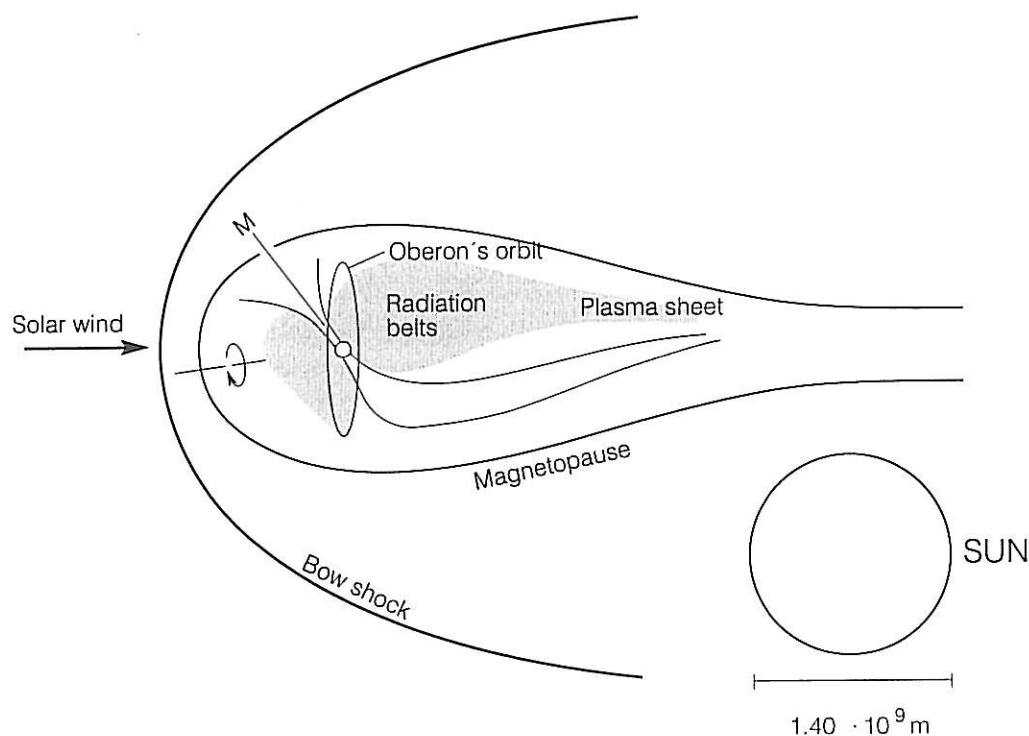


Figure 4.9.7. Sketch of the magnetosphere of Uranus (after Lanzerotti and Uberoi 1989).

summary of what the results from Voyager 2 revealed about it has been given by Krimigis (1990).

The magnetopause was passed at about 26 planetary radii. On the basis of magnetometer data from the only passage through the magnetosphere one has made a theoretical model of the magnetic field (Ness *et al.* 1989). It deviates strongly from a centered dipole and the field strength at the planetary surface varies within wide limits, about 20–150 μT . Its polarity relative to the rotation axis is opposite to that of the Earth, and the angle between them is about 47° . It has been possible to confirm that Neptune has radiation belts of both electrons and protons (Krimigis *et al.* 1989). The plasma in the magnetosphere is dominated by protons, but contains also ions of atomic nitrogen (Belcher *et al.* 1989). According to Krimigis (1990) only a small part of the plasma comes from the solar wind. (The atomic hydrogen ions, for example, may come from the upper atmosphere of Triton.)

Comets

The comets have no intrinsic magnetic field. On the other hand, magnetic structures are formed as a consequence of the “frozen-in” magnetic field lines (cf. Chapter 6 and Fig. 4.9.8.) being “hung up” on the comets coma (from Greek *come* = hair) – the diffuse gaseous envelope surrounding the nucleus of the comet.

During its recent passage (March 1986) the famous Halley’s comet was observed with no less than six space probes. Four of them, the Soviet Vega 1 and Vega 2, the Japanese Suisei and the European Giotto, passed the sunlit side at distances from 600 to 15 000 km. The Japanese probe Sakigake and the international ICE

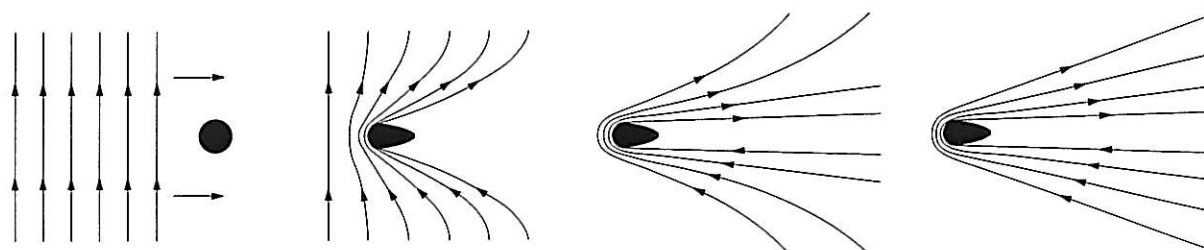


Figure 4.9.8. Formation of a cometary tail through the “frozen-in” magnetic field lines of the solar wind being “hung up” on the comets coma (after Alfvén, 1957).

(International Cometary Explorer) gathered data at greater distances upstream of the comet. Giotto, which was equipped with protection against dust particles passed as close as 600 km from the comet with a speed of 68.4 km/s and penetrated inside the contact surface (“ionopause”) between the solar wind and the cometary coma. One also found a bow shock in front of the ionopause. Fig. 4.9.9 shows a sketch of the environment of Halley’s comet based on the measurements of the space probes, especially Giotto’s, and Fig. 4.9.10 shows the cometary nucleus itself photographed by Giotto.

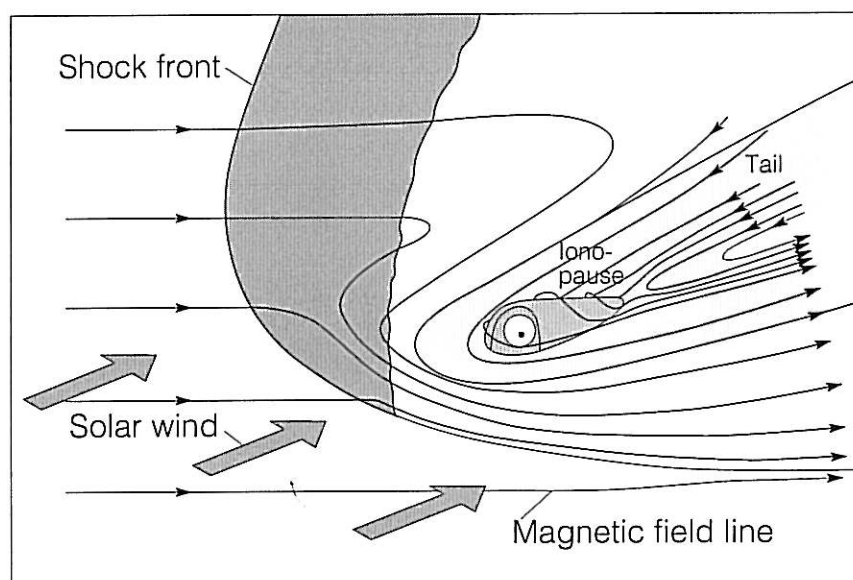


Figure 4.9.9. The “magnetosphere” of the comet Halley (Balsiger *et al.* 1988).

The comets are believed to come from a giant cloud (*Oort’s cloud*) of fragments from the formation of the solar system (see for example Weissman 1987). The Oort cloud is considered to have a radius about 50 000 times that of the Earth’s orbit around the Sun. When a clump of matter in this cloud is perturbed in its orbit so that it passes close to the Sun, it becomes a comet.



Figure 4.9.10. Photo, from the space probe Giotto, of the nucleus of Halley's comet (after Balsiger *et al.* 1988).

The Heliosphere

Just as plasma pressure and centrifugal force extends Jupiter's magnetosphere, it is supposed that the solar wind causes the formation of a magnetosphere around the Sun and the solar system – the *heliosphere* (Fig. 4.9.11). Its boundary is located where the dynamic pressure of the solar wind becomes too weak to push away the interstellar medium and its magnetic field. It is expected that at that location a standing shock wave is formed, at which the solar wind velocity abruptly decreases from supersonic to subsonic. In the interaction between the solar wind and the interstellar medium it is also believed that the critical velocity (see Section 6.4) plays a role.

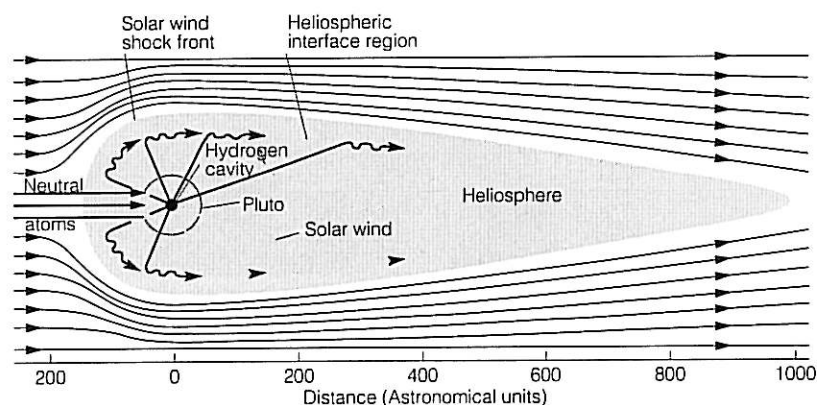


Figure 4.9.11. Sketch of the heliosphere (after Paresce and Bowyer 1986).

It is not known where the outer border of the heliosphere – the heliopause – is located, but it is usually expected to be at a distance of about 100 astronomic units (*i.e.* about $1.5 \cdot 10^{13}$ m). If the Voyager probes remain functional long enough, they may provide the answer.

Pulsars

The *pulsars* (cf. Chapter 7) are believed to be neutron stars with extremely strong magnetic fields (10 T or more at the surface) and therefore equipped with magnetospheres, which are the subject of intense theorizing. In pulsars, however, matter is in states so extreme, and so vastly different from anything we know experimentally, that any theory must be viewed with great caution. Experiences from the exploration of the Earth's own magnetosphere should inspire caution not to confuse theoretical models with reality.

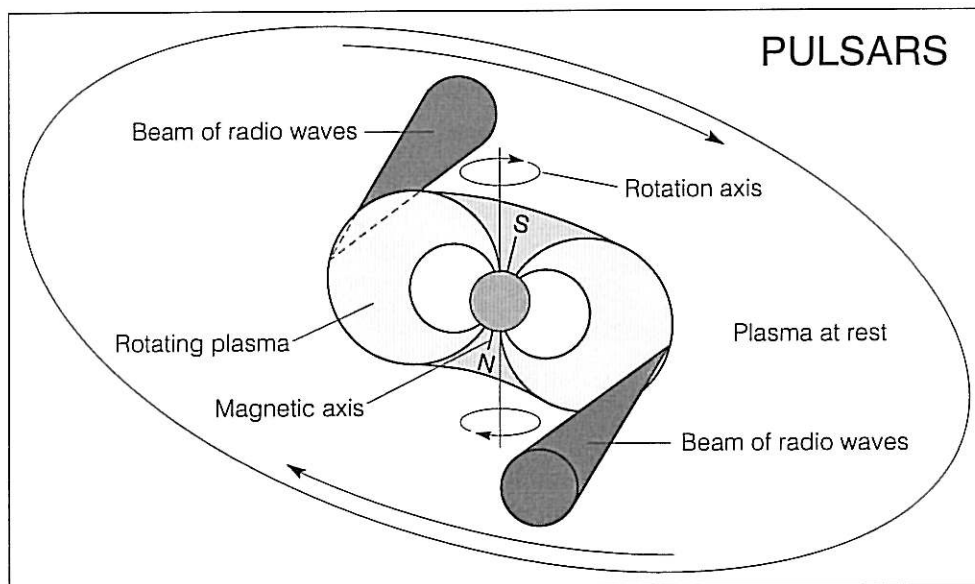


Figure 4.9.12. Pulsar atmosphere as usually envisaged

5. The Sun

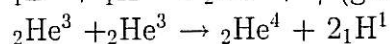
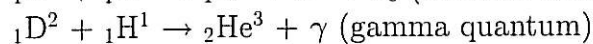
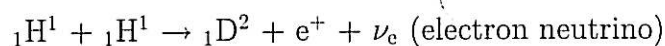
The Sun, which by its energy production is a prerequisite for the existence of life on Earth, is a comparatively small star of spectral class G (see Chapter 7). Like all other stars it consists of matter in the plasma state. Because the Sun is so close, it is possible to study in detail the manifold of physical phenomena that take place in this plasma. Since these have important consequences for space activities around the Earth we will here give a brief description of basic facts about the Sun and the *solar activity*.

Size, Mass and Power

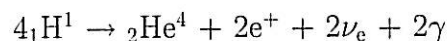
The Sun's *diameter*, $1.39 \cdot 10^9$ m, is about 109 times that of the Earth. Its *mass*, which is $2 \cdot 10^{30}$ kg, corresponds to about 333 000 Earth masses. Thus the average density is about 1.4 g/cm^3 (the corresponding number for the Earth is 5.5 g/cm^3). The total power released by the sun is $4 \cdot 10^{26}$ W.

Thermonuclear Fusion Reactions

The source of the Sun's energy is nuclear reactions in its interior, namely thermonuclear fusion reactions, in which hydrogen nuclei combine to form helium nucle The reaction rates are strongly temperature dependent and take place mainly inside about 0.25 solar radii, where the temperature is highest. The temperature at the center of the Sun is considered to be about 15 million K. The transformation of hydrogen to helium takes place by way of cyclic nuclear reactions, of which the most important is considered to be the so-called *proton cycle*:



The resulting reaction is thus a synthesis of helium from hydrogen:



This reaction implies transformation of mass into energy. The energy is found both as kinetic energy of the material reaction products (the helium nuclei and the positrons), and in the neutrinos and the gamma radiation. According to Einstein's well-known relation $E = mc^2$, the release of the $4 \cdot 10^{26}$ W mentioned above implies that the Sun uses up 4 million tons of its mass every second. This is, however, negligible in relation to the total mass of the Sun. (On the other hand, it is not entirely negligible in relation to the hydrogen reserve in the inner parts of the Sun where the reactions take place, and this plays a role in the evolution of the Sun on the time scale of billions of years.)

In spite of the impressive total power of the sun, the power release per unit mass is very moderate. Even in the central parts it is only of the order of mW/kg. This can be compared with the power released in the human body, which is of the order of W/kg. The reason is that fusion reactions between ordinary hydrogen-nuclei take place slowly, even at high temperatures. They are therefore useless for energy production on the Earth, where the hope is instead attached to heavy hydrogen, *i.e.* deuterium and tritium, the fusion reactions of which occur much more rapidly. The reason why the fusion of light hydrogen functions so well heating in the sun is simply the sun's enormous size, which both allows confinement of the plasma by means of its own gravitation, and causes the integrated power to be very large.

Energy Transport

From the source in the solar interior the energy is transported outward through a combination of several processes. Ordinary *heat conduction* plays a certain role, but is not dominating at any level. Nearest to the source *radiative transport*, dominates, *i.e.* emission and absorption of photons (mainly in the X ray and gamma ray wavelength range). In the upper layers of the Sun *convective transport* dominates, in which heat is transported by violent convective motions (which at the Sun's surface is visibly manifested in the form of granulation).

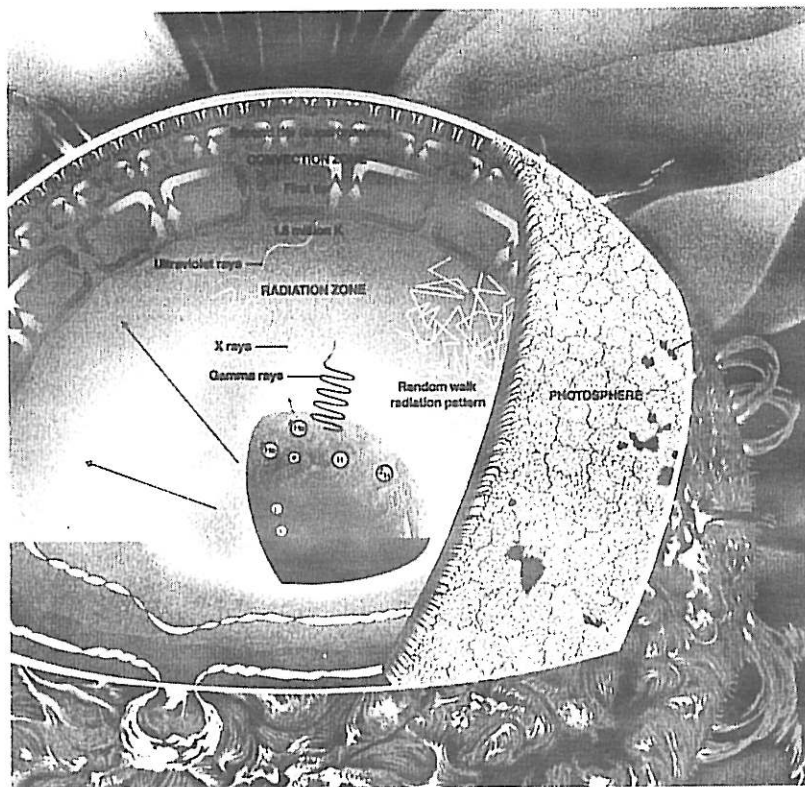


Figure 5.1. The interior of the Sun (after Friedman 1986).

Temperature and density decrease gradually outwards and at half the solar radius are, respectively, 3 million degrees and 1 g/cm^3 . A density comparable to the air density at the Earth's surface is reached at about 0.9 solar radii, where the temperature is a couple of hundred thousand degrees. The visible solar surface represents the level where the mean free path of the photons becomes so large that they can reach directly into space. At this level the temperature drops rapidly because of radiative losses. In this region, the *photosphere*, the temperature falls from about 6 000 K to about 4 600 K.

The Solar Atmosphere

The *photosphere* ("light" sphere), whose thickness is some hundred kilometers, constitutes the lowest layer of the *solar atmosphere*. It has a granular structure (see Fig. 5.2), which is called *granulation* and reflects a cellular convection with convection cells of the order of magnitude of 300 – 1 800 km and a life time of 1 – 10 minutes. The temperature difference between light and dark photospheric areas is about 100 K.

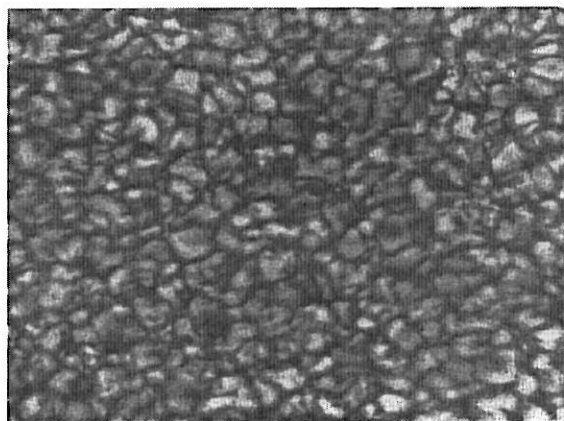


Figure 5.2. The granulation of the Sun's photosphere (Wyller 1986).

At the upper boundary of the photosphere the lowest temperature of the solar atmosphere is found, see Fig. 5.3.

Immediately above the photosphere is the *chromosphere* ("colour" sphere), which has its name from the fact that during solar eclipse it appears reddish. It is in this region, about 10 000 – 15 000 km deep, that the solar spectrum obtains its characteristic absorption lines, the *Fraunhofer lines*. In the chromosphere the temperature rises again so as to reach about 1 million degrees at the border of the corona. The apparently strange fact, that the temperature rises outward from the Sun – in the *same* direction as the energy transport – is considered to arise because turbulent motions in the lower solar atmosphere generate waves – sound waves and Alfvén

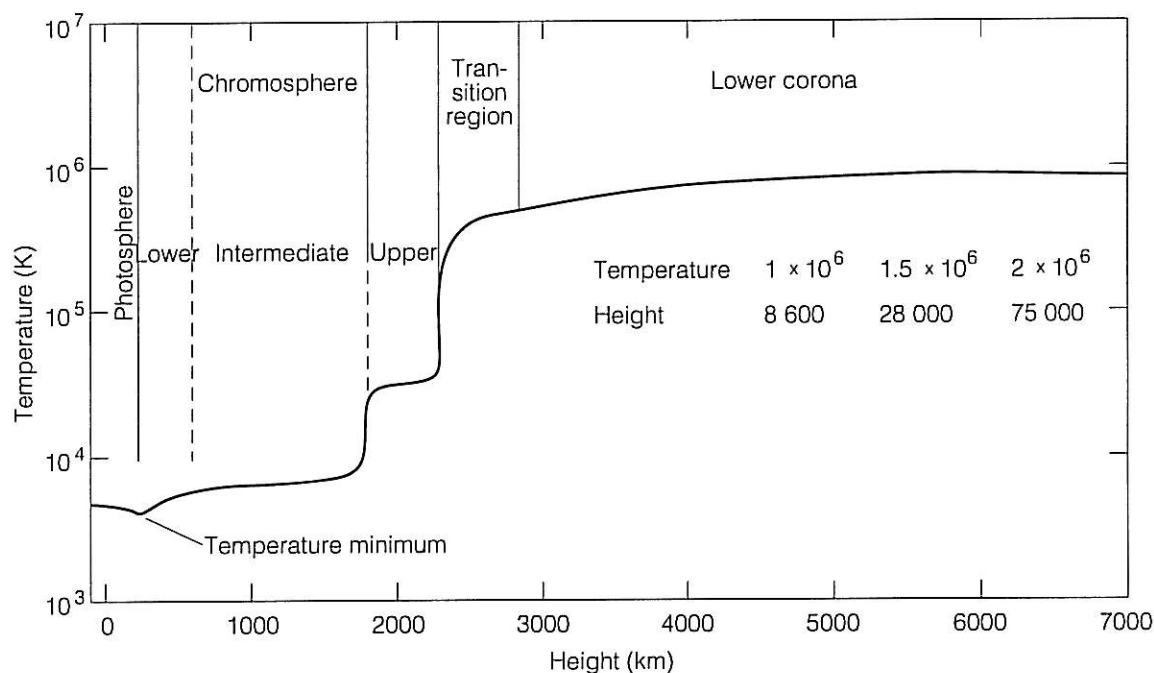


Figure 5.3. Distribution of average temperature in the solar atmosphere (Athay 1976).

waves – which at higher levels transform to shock waves and release their energy in the form of heat in the ambient medium.

The *Corona* (“crown”) is much fainter than the chromosphere and can only be observed during eclipse conditions or by means of a special instrument, the *coronagraph*, which efficiently shields the light from the photosphere. In the corona the temperature rises further to more than 2 million degrees. (Since there is no local thermodynamic equilibrium, we are here talking about kinetic temperature, cf. Section 2.2.). The high temperatures are derived from spectroscopic measurements. Initially it was difficult to identify the spectral lines of the corona, and one was inclined to believe that they derived from an element unknown on Earth, which was given the name *coronium*, until the Swedish astronomer Edlén in 1941 showed that they came from known elements in highly ionized states. (For example the so-called green corona line with the wavelength 5303 Å comes from 13-fold ionized iron.)

The density of the corona, which at an altitude of one solar radius above the photosphere is about 10^{-18} g/cm³, decreases without sharp border to the low values characterizing the solar wind, 10^{-24} – 10^{-23} g/cm³. The corona usually exhibits a pronounced ray structure, which probably is related to an inhomogeneous magnetic field. By means of radio astronomy methods this ray structure can be traced tens of solar radii into space. The structure of the corona varies in a characteristic way with the solar cycle (see below). Fig. 5.4 shows the typical appearance of the corona at the sunspot maximum and the sunspot minimum.

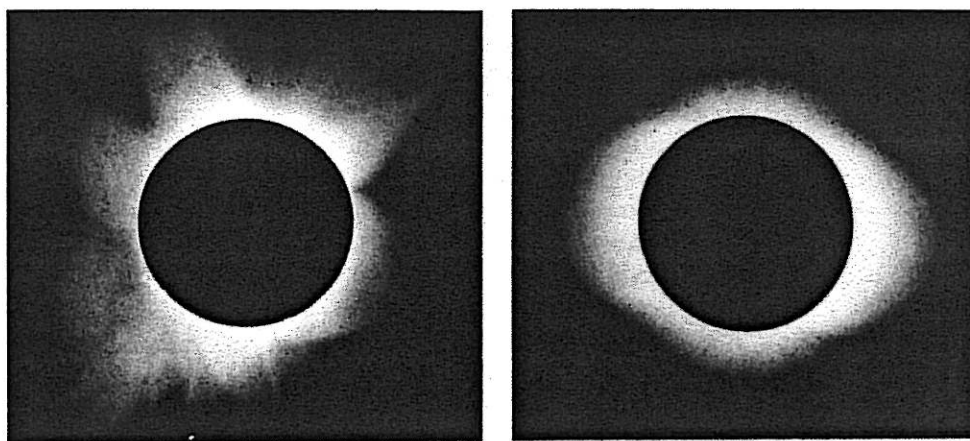


Figure 5.4. The solar corona at (a) sunspot maximum, (b) sunspot minimum (Menzel 1959).

The Sun's Magnetic Field

Magnetic fields on the Sun were first observed by Hale (1908) in sunspots (cf. below) and the Sun's general magnetic field was discovered a decade later (Hale *et al.* 1918) by means of the *Zeeman effect*. From measurements with the Zeeman effect one finds that average magnetic field of the Sun is weak, 1 – 3 Gauss. It was however pointed out by Hannes Alfvén (Alfvén and Fälthammar 1963), that the Sun's magnetic field can be largely hidden in thin magnetic “flux ropes”, with dimensions less than the resolution of the Zeeman effect measurements. It has later been found that measurements with higher resolution have revealed strong local magnetic fields, but measurements with even higher resolution are likely to be necessary before one can with certainty determine the Sun's total magnetic flux.

The *polarity* of the general magnetic field is opposite in the northern and in the southern hemisphere of the Sun, and this polarity changes in a 22-year cycle (cf. below under solar activity).

Rotation of the Sun

The Sun rotates with a period of about 27 days. It does not, however, rotate like a rigid body but exhibits a *differential rotation* with a maximum angular velocity at the equator. The rotational period at a certain latitude, λ , can approximately be calculated from the expression

$$25/(1 - 0.19\sin^2\lambda) \quad (5.1)$$

(Alfvén 1950, p. 111). An explanation, also proposed by Hannes Alfvén, of the differential rotation is that the early Sun lost angular momentum to the surrounding interplanetary plasma by means of electric currents along magnetic field lines from the Sun. Since the Sun's polar areas are magnetically coupled with larger and more distant volumes of interplanetary plasma, they should also have been subjected to the greatest loss of angular momentum.

Solar Activity

In spite of the fact that the energy supply from the interior of the Sun is continuous, there occur on the solar surface and in its vicinity a number of phenomena, termed *solar activity*, which is very unevenly distributed in space and exhibits strong time variation. These disturbances lead to emission of both electromagnetic and corpuscular radiation, which must be taken into account because of their potentially dangerous consequences for both men and instruments in orbit around the Earth. Some of the most important kinds of solar activity will be briefly described below.

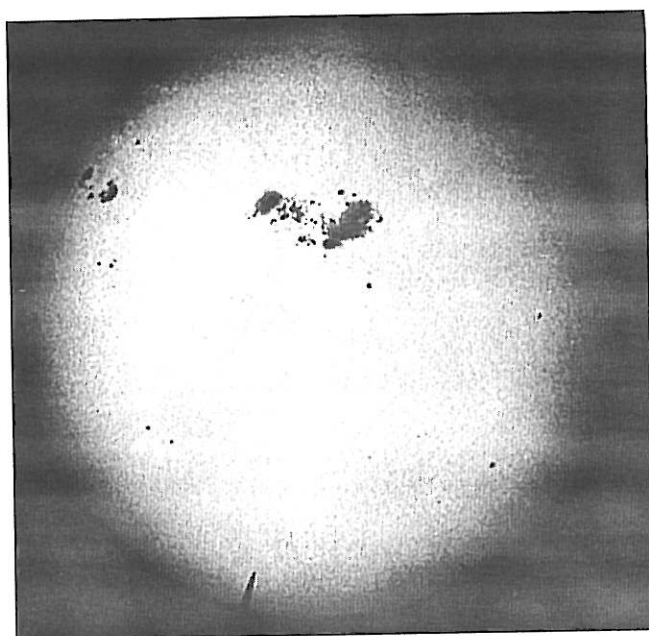


Figure 5.5. Sunspots

Sunspots are areas which have a low temperature compared to the surrounding photosphere (about 1 000 – 1 500 K lower) and therefore appear darker. (Fig. 5.5). The size of the sunspots vary from 100 km to about 100 000 km (the ones larger than 50 000 km can be observed with a naked eye). Often the sunspots consist of an inner part, the *umbra* (from lat. *umbra* = shadow) and an outer, somewhat brighter ring, the *penumbra* (from lat. *paene* = nearly), see Fig. 5.6. In many cases, however, penumbra is absent, and this is especially true for small sunspots. In the penumbra outward motions with a velocity of a few km/s are common – the so-called *Evershed-effect*. These motions appear to be part of a circulation which closes through upward motions outside the penumbra and at lower level back into the umbra. (There are also small penumbraless spots called “pores”).

The sunspots usually appear in pairs (sometimes in more complicated patterns) and are penetrated by strong magnetic fields – up to 0.4 T have been measured (by means of the Zeeman effect on the spectral lines of the Sun's light). The strong magnetic fields are the cause of the lower temperature of the sunspots because they impede the heat transport. The two spots in a sunspot pair have opposite magnetic

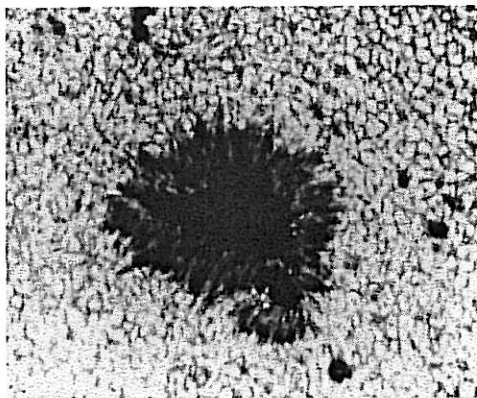


Figure 5.6. Close-up of a sunspot

polarity. The lifetime of the sunspots is usually a couple of weeks. (The above-mentioned pores, on the other hand, live only 10 – 60 minutes.)

Plages are areas in the chromosphere with increased emission as a consequence of higher temperature and density than the surroundings. (The explanation of the name is that early observers associated these brighter areas with white beaches.) The area is initially small, but grows and can reach a couple of per cent of the solar surface. These areas usually also have stronger magnetic fields, 200 – 500 Gauss (0.02 – 0.05 T) than the 1 – 3 Gauss that are typical of their surroundings. Within the plage area a fine structure of spots and extended shapes with a size of a few hundred km can be distinguished.

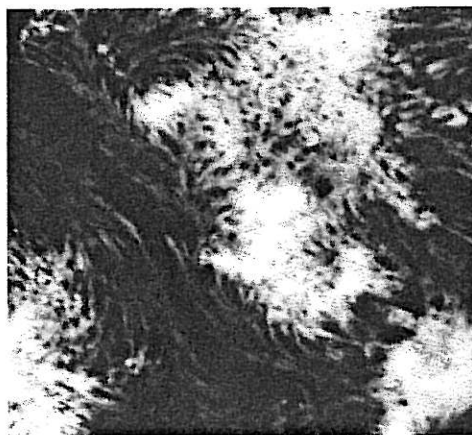


Figure 5.7. Plages (Bruzek and Durant 1977).

Spicules are filamentary plasma structures which penetrate through the chromosphere at a speed of about 30 km/s and which, when seen at the edge of the Sun, look like a forest. Individual spicules are about 1 000 km wide and up to 10 000

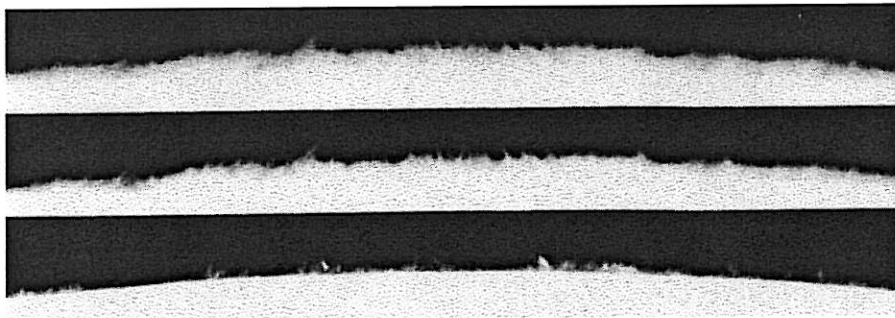


Figure 5.8. Spicules (*Handbook of Geophysics and the Space Environment* 1985).

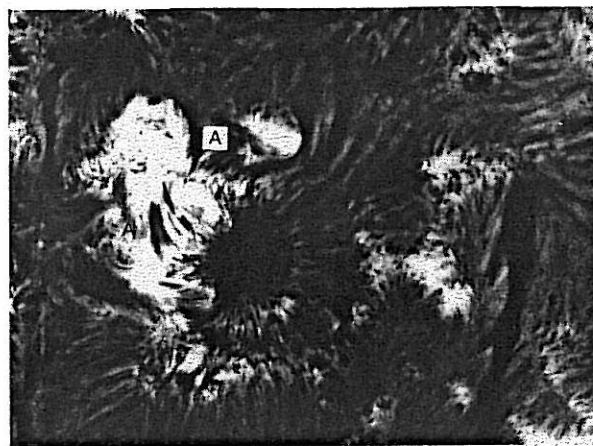


Figure 5.9. Fibrilles (Bruzek and Durant 1977).

km high. They have a temperature of 10 000 – 20 000 K and a lifetime of about 10 minutes.

Fibrilles are horizontal chromospheric plasma filaments, 1 000 – 2 000 km wide and about 10 000 km long, which often extend between the sunspots in a sunspot pair much like iron filings between magnetic poles. Their life time is 10 – 20 minutes.

Faculae are structures with enhanced emission of white light. They occur in the photosphere often together with chromospheric plages and precursors of sunspots. The faculae are long-lived and often remain even when the sunspot which they have preceded is already gone.

Prominences are a very prominent phenomenon on the Sun. They are veils of plasma which extend outward from the Sun's surface or arise as arc-like structures above it, see Fig. 5.11. Their length is often 200 000 km while width and height are of the order of 6 000 and 40 000 km respectively. From an observational point of view they appear as filaments – *dark* if seen against the photosphere, *bright* if observed at the Sun's rim with the dark space as a background. The prominences appear to be ejected from the Sun's surface, but recordings show that in quiet prominences the motion of the matter is downward along the magnetically controlled structures. One distinguishes between *quiet* and *eruptive* prominences. The latter are characterized,

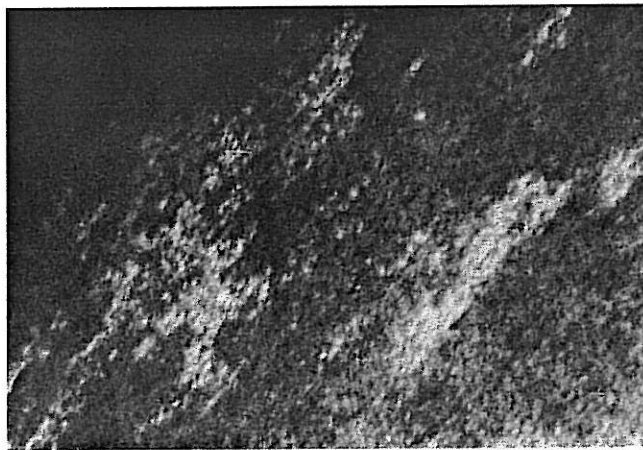


Figure 5.10. Faculae (Bruzek and Durant 1977).

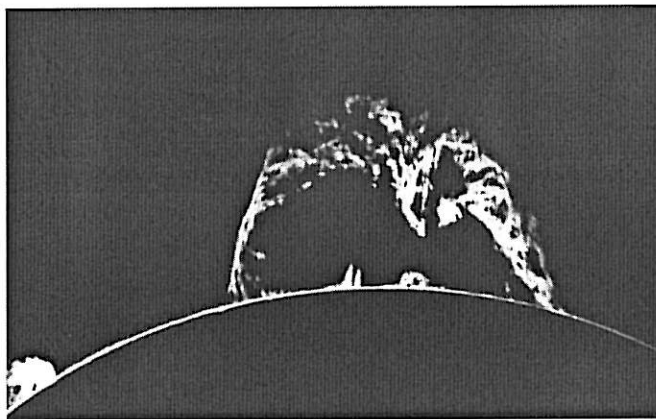


Figure 5.11. Prominences (Bruzek and Durant 1977).

as the name implies, by outward motions which can reach a speed of hundreds of km/s.

The *solar eruption* or the *flare* is the most important kind of solar activity from the point of view of space activity. It is a sudden local intensification of radiation from the Sun. It often occurs in connection with pairs or groups of sun spots. Fig. 5.12 shows an example of a solar flare. A solar flare can, in the time span of some minutes, release energy of the order of $10^{22} - 10^{25}$ J. While the emission of visible light increases only a few per cent, and which does not substantially influence the total irradiance of the Sun, the increase of shortwave radiation is dramatic. In the X-ray wavelength range emission can increase ten thousandfold.

In addition to the effects that a flare has on the ionosphere and which has been described in Section 3.3, it drastically influences the space environment in the Earth's vicinity, among other things through the *energetic particles*, mainly *protons*, which it emits, and which constitute a danger to both humans and electronics in spacecrafts. Solar flares are classified on the basis of their surface, measured in millionths of the surface of the solar hemisphere, in five classes of importance. These are defined in the table below.

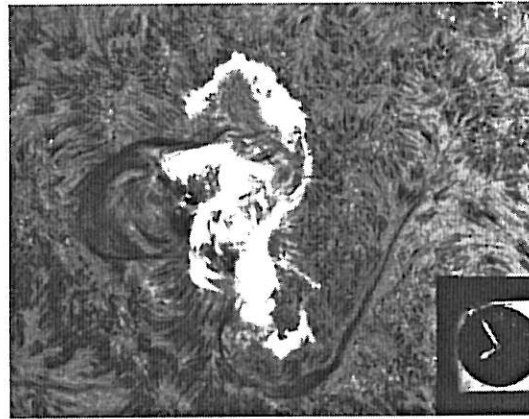


Figure 5.12. Solar flare (Big Bear Solar Observatory, August 7, 1972).

Areal importance	Area in "millionths"	Square degrees
S	< 100	< 2.0
1	100 - 250	2.1 - 5.1
2	250 - 600	5.2 - 12.4
3	600 - 1200	12.5 - 24.7
4	> 1200	> 24.7

Table 5.1. Classification of solar flares. The area is expressed in millionths of the hemisphere. (Bruzek and Durant 1977, p. 84).

The mechanism of solar flares is not yet definitely understood. It is, however, generally believed that what happens is a sudden transformation of magnetically stored energy to kinetic energy of charged particles, which in turn causes the electromagnetic radiation. A particular difficulty is that the plasma of the solar atmosphere is a good electric conductor, which makes it very hard to rapidly release magnetic energy.

A possible explanation, proposed by Hannes Alfvén and his collaborators at the Royal Institute of Technology in Stockholm, is that in the current circuit which is the source of the magnetic field, there is created one or more *electric double layers* with a very high voltage. When the inductance of the circuit forces the current to flow through this voltage drop, all the current carriers are accelerated to an energy corresponding to the voltage drop (which they subsequently share through collisions with other particles and with the electromagnetic radiation field). For a current circuit of the size of a sun spot group the inductance has been estimated at $10 - 1\,000\text{ H}$ and the magnetic energy at 10^{25} J . This energy can be released in the time span characteristic of the solar flare if the voltage drop of the double layers

is $10^{10} - 10^{11}$ V. This also turns out to correspond well to the energy of the most energetic particles in a solar flare (Carlqvist 1969, 1986).

The Solar Cycle

The occurrence of sunspots exhibits a periodic variation, the *solar cycle*, with a period of about 11 years. From a minimum the activity rises in about 4 years to a maximum and subsequently decreases during about 7 years to a new minimum. (The length of the solar cycle can vary between 7 and 13 years.) Individual solar cycles are counted from minimum to minimum. The present solar cycle which began in September 1988 has been given the number 22.

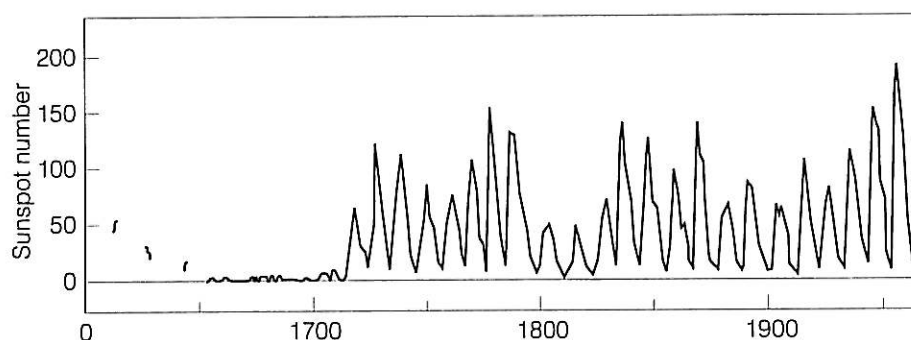


Figure 5.13. The longtime variation of the sunspot number (Robinson 1987).

If the occurrence of sunspots is presented in a diagram with time on the horizontal axis and latitude on the vertical axis, the diagram gets an appearance which is illustrated in Fig. 5.14, and therefore is called *butterfly diagram*. The diagram shows that at the beginning of a new solar cycle the sun spots appear at about 40° latitude. Subsequently they appear at lower and lower latitudes until at the end of the cycle they appear close to the Equator.

Different solar cycles can be very different. In historical time very large variations have occurred. There are documents that indicate that during a period of 70 years, 1645–1715, almost no solar activity at all occurred (the *Maunder-minimum*), cf. Fig. 5.13. As a measure of solar activity it is customary to use a weighted combination of the number of sunspots and the number of sunspot groups. This “*sunspot number*” is published regularly in astronomical journals. The greatest sunspot activity registered until now occurred during the 1957–58 maximum (which coincided with the *International Geophysical Year*) with a sunspot number exceeding 200. During a sunspot minimum sunspots may be entirely absent for several days.

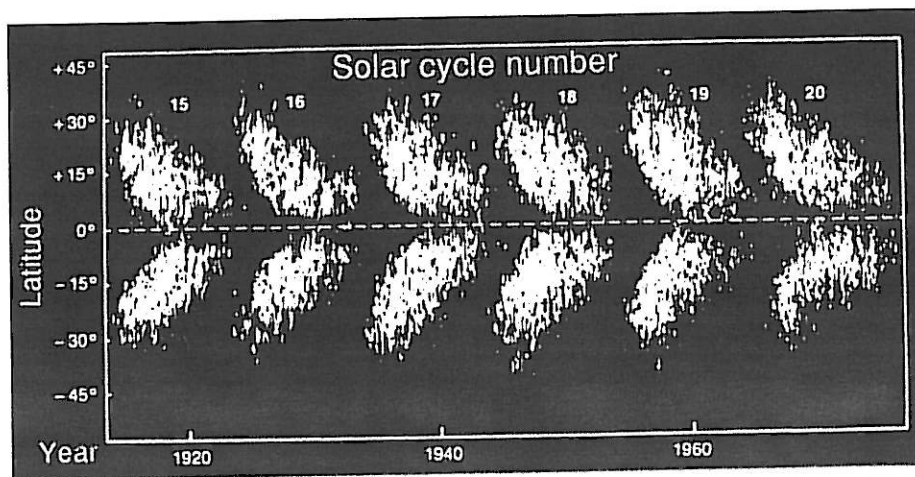


Figure 5.14. Butterfly diagram showing the distribution of sunspots during the solar cycles 15–20 (Robinson 1987).

The general solar magnetic field mentioned before has different polarity during subsequent solar cycles. If the magnetic field polarity is taken into account as well as the sun spot activity, one ends up with a periodicity of 22 years instead of 11-years. During the present solar cycle the polarity is negative in the northern hemisphere, which means that the direction of the magnetic field is *into* the Sun in the northern hemisphere.

6. Interplanetary Space

6.1. The Solar Wind

Interplanetary space is filled with a thin plasma, the *solar wind*, which continually streams out of the Sun and traverses the solar system at a speed of 300 – 900 km/s. The solar wind is the link between phenomena on the Sun and their effects on the environment of the Earth, especially magnetic storms and auroras.

The Comets as Solar Wind Probes

The first prediction of a continual outflow of plasma (“Korpuskularstrahlung”) from the Sun was made by Biermann on the basis of observations of cometary tails, see *e.g.* Biermann and Lüst (1963). A certain type of cometary tails, Type I, consist of ionized gases, *i.e.* plasma (mainly CO_2 , N_2^+ , CO_2^+ och CH^+). As distinguished from cometary tails of Type II, which consist of dust particles and have a curvature which reflects the orbital motion of the comet, the plasma tails are straight because coupling to the solar wind plasma causes them to rapidly accelerate to so high a radial velocity that the orbital velocity becomes negligible. Detailed study of inhomogeneities in the tails reveals accelerations radially from the Sun of up to 10 m/s^2 . It was in order to explain these rapid accelerations that it was found necessary to introduce the assumption of a plasma wind from the Sun, which tears with it the ionized but not the unionized cometary products.

A remaining problem, however, is how the gas can be ionized so rapidly. Calculations show that the Sun’s electromagnetic radiation is insufficient for this, and the particle radiation does little to ionize, since collisions with neutral gas particles are very few. A probable explanation is a plasma phenomenon called the critical velocity (cf. Section 6.4).

Origin of the Solar Wind

The plasma flow from the Sun may seem to be a simple consequence of the energy supply to the corona, which is reflected in its high temperature (over 2 million degrees). A remarkable fact is, however, that the flow becomes supersonic, both with respect to the ordinary sound speed and with respect to the other important signal

velocity, the Alfvén velocity (described in Section 1.3). It is well known from gas dynamics that in order to cause a streaming gas to reach supersonic speed, it is necessary to use a special device, called *i.e.* Laval–nozzle, where the cross-section of the flow has a minimum. It is at this minimum cross-section that the transition from subsonic to supersonic flow can take place. The solar corona is not equipped with a Laval–nozzle. It can, however, be shown (Section 6.2) that the Sun’s gravitational field adds to the equations of motion of the plasma a term that is formally similar to the one that describes the effects of the Laval–nozzle. The resulting equations have solutions that describe a transition from subsonic to supersonic flow at a distance of a couple of solar radii.

This does not, however, solve the entire problem concerning the acceleration of the solar wind. The elegant gas dynamic analogy needs to be amended by taking into account the fact that the solar wind like other cosmic plasmas, is inhomogeneous and that electrodynamic forces can play an important, maybe completely dominating, role (Carlqvist and Alfvén 1980).

Properties of the Solar Wind

On the basis of measurements by means of space probes the properties of the solar wind are since long well known. Typical values of density, velocity, particle flux, energy flux, temperature and magnetic field at the Earth’s orbit are summarized in Table 6.1.1.

Proton density	$8 \cdot 10^6 \text{ m}^{-3}$ (8 cm^{-3})
Solar wind velocity	$3.2 \cdot 10^5 \text{ m/s}$ (320 km/s)
Proton flux	$2.4 \cdot 10^{12} \text{ m}^{-2} \cdot \text{s}^{-1}$
Kinetic energy flux	$2.2 \cdot 10^{-4} \text{ W/m}^2$ (0.22 mW/m^2)
Proton temperature	$4 \cdot 10^4 \text{ K}$
Electron temperature	10^5 K
Magnetic field strength	5 nT

Table 6.1.1. Typical solar wind parameters (Fälthammar 1973a).

As can be seen from the table, the electron and proton temperatures are different, which is a common occurrence in dilute cosmical plasmas, because the collisions are few and collisions between particles of different mass are less efficient for exchange of momentum than collisions between particles of equal mass.

Energy density of massmotion	1
Thermal energy density of protons	1/120
Thermal energy density of electrons	1/50
Magnetic energy density	1/70

Table 6.1.2. Relative energy densities in the solar wind at the orbit of the Earth (Fälthammar 1973).

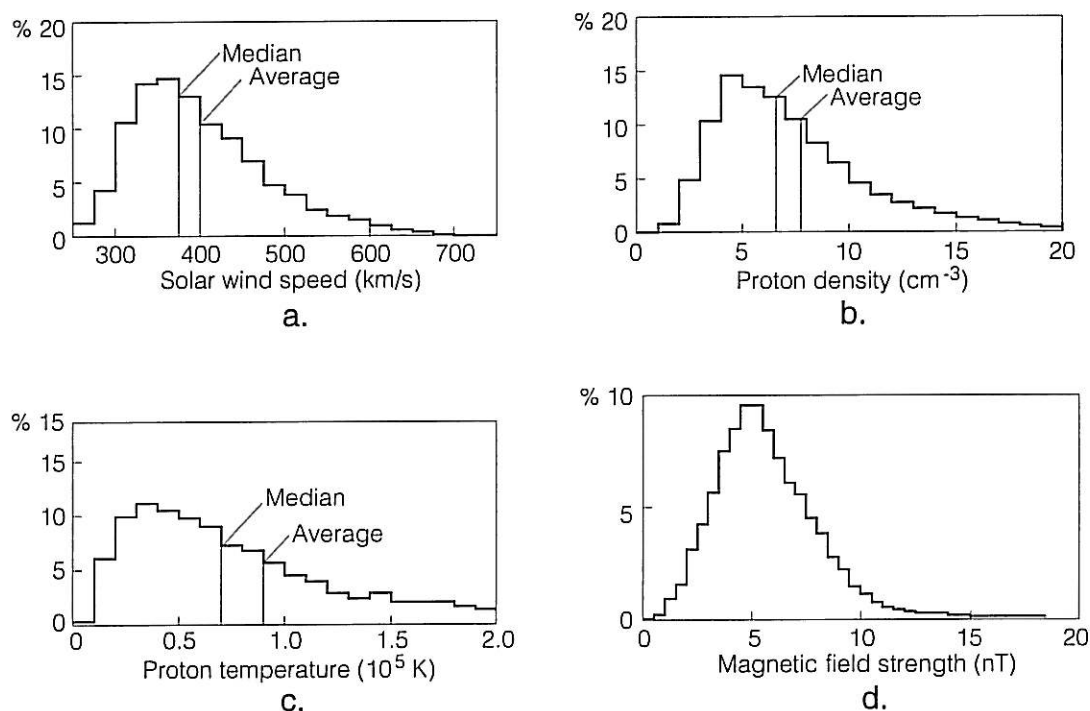


Figure 6.1.1. Histogram of the occurrence frequency of different values of the solar wind (a) velocity, (b) density, (c) proton temperature and (d) magnetic field strength (Hundhausen et al. 1970 (a-c), Ness 1969).

From the values in Table 6.1.1 one can calculate the energy density relations in the solar wind. These are summarized in Table 6.1.2.

An important property of the solar wind, that is not obvious from the table is its great variability. This can instead be appreciated from Fig. 6.1.1, which shows the relative occurrence of different values of the main physical parameters.

In terms of chemical composition the solar wind consists mainly of hydrogen (protons). The second largest constituent is helium (in the form of He^{2+} , *i.e.* alpha particles). The helium content is usually about 5 % but can occasionally rise to more than 20 %. The variations reflect still unknown chemical separation processes in the solar plasma, possibly related to the ones discovered in the Earth's ionosphere-magnetosphere system (cf. Section 1.1).

The Interplanetary Magnetic Field

If the interplanetary magnetic field derived only from currents in the Sun, it would rapidly decrease and have a completely negligible strength at the Earth's orbit. In reality it is maintained by electric currents in the solar wind itself and even at the Earth's orbit it has the strength of several nT, as shown in Table 6.1.1.

The large-scale properties of the interplanetary magnetic field can be simply derived from the famous law of magnetohydrodynamics, the law of frozen-in magnetic field lines (Section 1.3). This implies that in a medium of sufficiently large electrical conductivity the magnetic field and the material motion are mutually coupled, as if the magnetic fields were tied to the medium ("frozen-in"). This condition which was illustrated in Fig. 1.3.2, implies, in rigorous formulation, that if two elements of the medium are on a common magnetic field line at some instant of time, they will be on a common field line at any other instant of time, however much the magnetic field is deformed in between the elements, provided that the condition for the frozen-in condition has been satisfied all the time along the part of the magnetic field line that connects the two plasma elements. The condition for freezing-in (which is well satisfied in the solar wind) and the derivation of the law is given in Section 6.3.

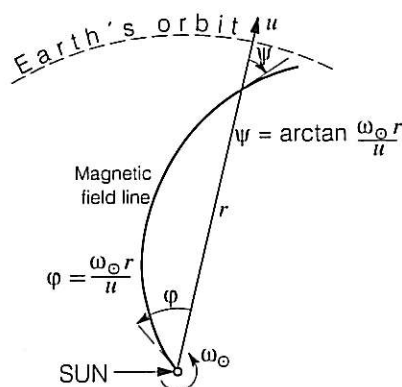


Figure 6.1.2. Spiral-shaped magnetic field line generated by the radial motion of the solar wind combined with the Sun's rotation

When the condition of frozen-in field lines is valid, the ratio of energy densities of the material motion and of the magnetic field determines which of them controls the other. If the magnetic field energy dominates, the structure and motion of the plasma is adjusted accordingly. (An example of this case is the ray structures in the aurora.) If the moving material has a much higher energy density than the magnetic field (as in the case of in the solar wind, see Table 6.1.2) the magnetic field follows the former essentially without resistance. However, if the energy densities are comparable there is a mutual interaction.

If this law is applied to the radial motion of the solar wind combined with the rotation of the Sun, it is found that the magnetic field must be pulled out into a spiral shape, as illustrated in Fig. 6.1.2.

The spiral angle ψ is determined by the condition that $\tan \psi = \omega_{\odot} r / u$, where ω_{\odot} is the angular velocity of the Sun and u is the radial velocity of the solar wind.

The resulting shape is an *Archimedes spiral*. At the Earth's orbit the angle ψ is about 45° .

For completeness it should be pointed out that the solar wind is not always precisely radial. *Fluctuations* in its direction of up to 10° occur. Furthermore, because of the Earth's orbital motion around the Sun with a speed of about 30 km/s, even a radially streaming solar wind appears to come from a direction which deviates somewhat toward the morning side. This *aberration* is about 5° and approximately inversely proportional to the solar wind velocity.

The spiral form just derived describes very well the *average* direction of the interplanetary magnetic field. However, the frozen-in magnetic field contains a great deal of irregularities, and the instantaneous magnetic field at a given point therefore varies over the full 360° angular range, as the polar histogram in Fig. 6.1.3 shows.

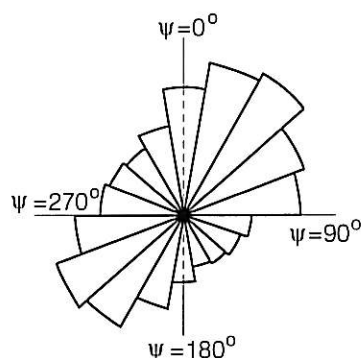


Figure 6.1.3. Polar histogram showing the relative occurrence frequency of different directions of the ecliptic component of the solar wind magnetic field. (Ness and Wilcox 1964).

The deviations from the average field also imply that interplanetary magnetic field often has a component transverse to the ecliptic plane, *i.e.* transverse to the orbital plane of the Earth. As mentioned in Section 4.4, this component plays a decisive role for the interaction of the solar wind with the Earth's magnetosphere.

The Interplanetary Current Sheet

Since, according to Maxwell's equations, the divergence of the magnetic field is zero, the magnetic flux out of the Sun must be precisely equal to the flux into the Sun. The interplanetary magnetic field must therefore be alternatively outward and inward directed. It has been found that there is a large-scale regularity in these changes of direction. The first measurements, which took place near the ecliptic plane, showed that there were "sectors" in which the magnetic field was preferentially inward respectively outward directed. As expected, this "sector structure" followed the Sun's rotation, since the polarity of the magnetic field line is determined by the conditions on the particular part of the Sun's surface where it has its footpoint. These sectors change slowly with time (on a time scale of weeks or months) and their number was sometimes two, sometimes four, and sometimes the pattern was more irregular.

Later it was found that it is not a matter of sectors. Instead the explanation is that the solar wind magnetic fields is not divided into sectors. It only appears that way when observation are limited to the ecliptic plane. The three-dimensional structure is such that opposite magnetic polarities prevail above and below a wavy boundary surface as shown in Fig. 6.1.4. A space probe in the ecliptic plane will therefore be located alternately in regions of outward and inward directed magnetic field, depending on the phase of the solar rotation. The changing number and extents of the various “sectors” are simply a result of the wavy boundary surface changing its shape.

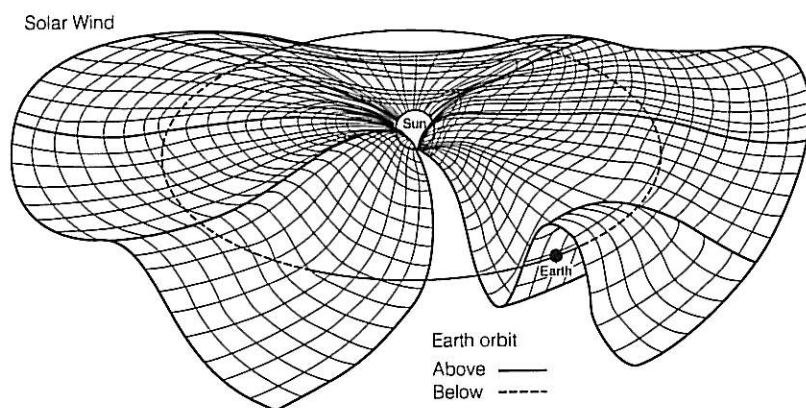


Figure 6.1.4. Sketch of the interplanetary current sheet (after Friedman 1986).

Since the magnetic field direction is opposite on the two sides of the boundary surface, there must, according to the Maxwell equation $\text{curl} \mathbf{B} = \mu_0 \mathbf{i}$, flow an electric current along this surface. (For a typical field strength of 5 nT the current density, integrated over the thickness of the boundary surface, is about 10^{-2} A/m.

Mach Number

Since the solar wind is electrically conducting and penetrated by a magnetic field, it can also propagate Alfvén waves (cf. Section 1.3). As a consequence, in the solar wind the Alfvén velocity is a signal velocity as important as the velocity of sound, and one has to consider two different Mach numbers, the ordinary *Mach number*, which is the ratio between the solar wind velocity and the local sound velocity, and the *Alfvén-Mach number*, which is the ratio between the solar wind velocity, u , and the local Alfvén velocity, V_A . From the values in Table 6.1.1 the following typical values can be calculated:

Sound velocity

$$V_S = \sqrt{\frac{\gamma k T_e}{m_p}} \approx 35 \text{ km/s} \quad (6.1.1)$$

Mach number

$$M = u/V_S \approx 9 \quad (6.1.2)$$

Alfvén velocity

$$V_A = \frac{B}{\sqrt{\mu_0 n m_p}} \approx 40 \text{ km/s} \quad (6.1.3)$$

Alfvén–Mach number

$$M_A = u/V_A \approx 8 \quad (6.1.4)$$

The Electric Field

As shown in Section 4.7, the electric field depends on the state of motion of the observer. We can use this to calculate the electric field in interplanetary space.

As the solar wind plasma is a good electric conductor, the electric field, \mathbf{E}' , in the reference frame of the plasma itself must be zero. If we denote the electric field in a fixed coordinate system by \mathbf{E} , the following relation should hold

$$\mathbf{E}' = \mathbf{E} + \mathbf{u} \times \mathbf{B} = 0$$

i.e.

$$\mathbf{E} = -\mathbf{u} \times \mathbf{B} \quad (6.1.5)$$

For typical values of the solar wind velocity and magnetic field the magnitude of \mathbf{E} is about 2 mV/m at the Earth's orbit. Its direction fluctuates with the changes of direction of the magnetic field (whereas the velocity vector \mathbf{u} always is nearly radial). As mentioned in Section 4.4, an electric field directed from the morning side toward the evening side of the Earth's magnetosphere corresponds to a strong interaction between the solar wind and the magnetosphere ("the rectifier" effect).

As in many other contexts, it is also true for the solar wind that the *energy density* of the electric field is very small. The ratio of electrical and magnetic energy is given by

$$\frac{\epsilon_0 E^2/2}{B^2/2\mu_0} = \epsilon_0 \mu_0 \left(\frac{E}{B} \right)^2 = \left(\frac{u}{c} \right)^2 \sin^2 \psi \quad (6.1.6)$$

where u is the solar wind velocity, c the velocity of light and ψ is the angle between \mathbf{B} and \mathbf{u} . (E/B is simply the solar wind velocity component transverse to the magnetic field.)

The solar wind can be considered as a cosmical MHD-generator. Over an area of the size of the cross section of the magnetosphere the solar wind electric field has a potential drop of hundreds of kilovolts, which represents the electromotive force of the solar wind. In the case of strong coupling (southward directed interplanetary magnetic field, see Section 4.4) electric currents flow to and from the Earth's environment and feed part of the energy from the generator to the interior of the magnetosphere, where it is dissipated partly in the form of energy of charged particles, *e.g.* in radiation belts and polar auroras, partly as Ohmic heating in the ionosphere of the auroral ovals.

6.2. Gas Dynamic Solar Wind Analogy

The remarkable fact that the solar wind is supersonic can, according to one theory, be explained by a gas dynamic analogy, where the Sun's gravitation plays the same role as the Laval-nozzle plays for example in a gas turbine. This theory is sketched below.

Start from the equations of continuity for *mass*, *momentum* and *energy*. For a Laval-nozzle, whose cross-section, S , varies with the axial coordinate, x , these equations become

$$\frac{d}{dx}(\rho_m u S(x)) = 0 \quad (6.2.1)$$

$$\rho_m u \frac{du}{dx} = -\frac{dp}{dx} \quad (6.2.2)$$

$$\frac{d}{dx}(p/\rho_m^\gamma) = 0 \quad (6.2.3)$$

where ρ_m , u and p are mass density, velocity and pressure, and $\gamma = C_p/C_v$ is the ratio between the specific heat at constant pressure and at constant temperature.

If one combines (6.2.1) – (6.2.3) and introduces the Mach number M , which is given by the expression

$$M = \frac{u}{V_s} = \frac{u}{\sqrt{\frac{\gamma p}{\rho_m}}} \quad (6.2.4)$$

one finds that the dependence of the Mach number on the axial coordinate, x , is described by the equation

$$\frac{M^2 - 1}{2M^2} \frac{dM^2}{dx} = \left(1 + \frac{\gamma - 1}{2} M^2\right) \frac{1}{S} \frac{dS}{dx} \quad (6.2.5)$$

As a transition from subsonic ($M < 1$) to supersonic ($M > 1$) flow requires that the factor $M^2 - 1$ on the left hand changes sign, the factor dS/dx on the right hand side must also change sign, which it does at the throat of the Laval-nozzle.

For the solar wind the continuity equations (6.2.1) – (6.2.3) have the form

$$\frac{d}{dr}(\rho_m u r^2) = 0 \quad (6.2.6)$$

$$\rho_m u \frac{du}{dr} = -\frac{dp}{dr} - \frac{G m_\odot \rho_m}{r^2} \quad (6.2.7)$$

$$\frac{d}{dr}(p/\rho_m^\gamma) = 0 \quad (6.2.8)$$

where m_\odot is the solar mass ($2 \cdot 10^{30}$ kg) and G the gravitational constant ($6.673 \cdot 10^{-11}$ Nm²/kg²).

The equations (6.2.6) – (6.2.8) can be combined to

$$\frac{d}{dr} \left(\frac{u^2}{2} + \frac{\gamma}{\gamma-1} \frac{p}{\rho_m} - \frac{Gm_\odot}{r} \right) = 0 \quad (6.2.9)$$

If also the Mach number defined by (6.2.4) is introduced here, a differential equation is obtained which describes the dependence of the Mach number on the radius r , namely

$$\frac{M^2 - 1}{2M^2} \frac{dM^2}{dr} = \left(1 + \frac{\gamma-1}{2} M^2 \right) F(r) \quad (6.2.10)$$

where the factor $F(r)$ is given by

$$F(r) = \frac{\frac{2}{r} - \frac{5-3\gamma}{\gamma-1} \frac{Gm_\odot}{2r^2 E_0}}{\frac{1+Gm_\odot}{2r E_0}} \quad (6.2.11)$$

and E_0 is a constant (equal to the quantity in parenthesis in (6.2.9)) that represents the energy per unit of mass (Holzer and Axford 1970).

Obviously the equations (6.2.5) and (6.2.10) are entirely analogous, and the factor $F(r)$ in the case of solar wind corresponds to the factor $S^{-1}(dS/dx)$ which is in the case of the Laval-nozzle. Without the gravitational term the factor $F(r)$ would not be able to change signs, and transition from subsonic to supersonic flow would be impossible. With the gravitational term present, such a change of signs takes place, and the equation (6.2.10) has solutions where M goes from $M < 1$ to $M > 1$ (Holzer and Axford 1970). Examples of solutions of (6.2.10) are shown in Fig. 6.2.1.

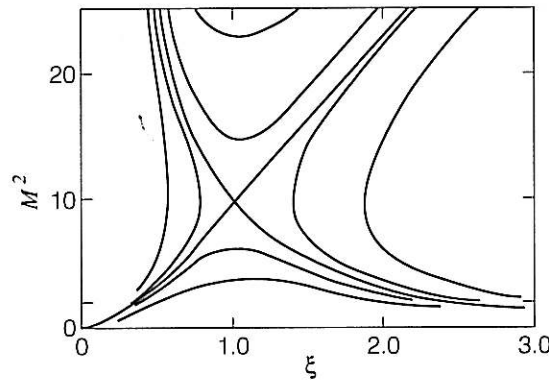


Figure 6.2.1. Graphic representation of solutions to (6.2.10), given as the relation between the square of the Mach number and the dimensionless radius, ξ (Holzer and Axford 1970).

6.3 Frozen-in Magnetic Field Lines

In connection with the discovery of the magnetohydrodynamic waves Alfvén (1940) noted that the motion of the matter and the deformation of the magnetic field were coupled in such a way that the magnetic field lines appeared to “follow” the motion of the matter in the way illustrated in Fig. 1.3.2. He gave this behaviour the name “*frozen-in magnetic field lines*”. Below will be given a general derivation of the law and of its conditions of validity.

Let \mathbf{B} be the magnetic field and \mathbf{v} the velocity field with respect to a fixed coordinate system, and \mathbf{E} the electric field in the system. Denote by \mathbf{E}' the electric field in a (local) coordinate system, following the motion of the medium. As the electric field that enters in Ohm's law is the field in the rest frame of the medium, equation (4.7.3) implies

$$\mathbf{i} = \sigma \mathbf{E}' = \sigma(\mathbf{E} + \mathbf{v} \times \mathbf{B}) \quad (6.3.1)$$

where \mathbf{i} is the current density and σ the conductivity. If this relation is combined with the Maxwell equations

$$\text{rot } \mathbf{B} = \mu_0 \mathbf{i} \quad (6.3.2)$$

$$\text{rot } \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (6.3.3)$$

where the displacement term is neglected in (6.3.2), one finds

$$\begin{aligned} \frac{\partial \mathbf{B}}{\partial t} &= -\left(\frac{1}{\mu_0 \sigma} \text{rot rot } \mathbf{B} - \text{rot}(\mathbf{v} \times \mathbf{B}) \right) \\ &= \text{rot}(\mathbf{v} \times \mathbf{B}) + \frac{1}{\mu_0 \sigma} \text{div grad } \mathbf{B} \end{aligned} \quad (6.3.4)$$

If l_c is a characteristic length and v_c a characteristic velocity, the relative magnitude between the second last and the last term in (6.3.4) can be written

$$R_m = \mu_0 \sigma l_c v_c \quad (6.3.5)$$

This ratio is usually called the *magnetic Reynold number*. If $R_m \ll 1$, (6.3.4) becomes

$$\frac{\partial \mathbf{B}}{\partial t} = \frac{1}{\mu_0 \sigma} \text{div grad } \mathbf{B} \quad (6.3.6)$$

which has the form of a *diffusion equation*, with the *diffusion constant*

$$D = \frac{1}{\mu_0 \sigma} \quad (6.3.7)$$

The case that interests us here is the opposite, when

$$R_m \gg 1 \quad (6.3.8)$$

so that (6.3.4) is reduced to

$$\frac{\partial \mathbf{B}}{\partial t} = \text{rot}(\mathbf{v} \times \mathbf{B}) \quad (6.3.9)$$

which has the following interesting consequence.

Consider the magnetic flux Φ through a surface with the edge contour C which follows the motion of the plasma. Since the motion of the plasma is in general different at different points along C , this contour becomes deformed during the motion. The total change per unit time of magnetic flux through the surface is therefore the sum of one part which depends on the time derivative of the magnetic field at points inside the contour and another which depends on the fact that the contour moves so as to cut across magnetic field lines. The latter contribution is given locally by the product $\mathbf{v} \times \mathbf{B}$. The total change of flux per unit time therefore becomes (see Fig. 6.3.1)

$$\begin{aligned} \frac{d\Phi}{dt} &= \iint \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} + \oint (\mathbf{v} \times d\mathbf{s}) \cdot \mathbf{B} \\ &= \iint \frac{\partial \mathbf{B}}{\partial t} \cdot d\mathbf{S} - \oint (\mathbf{v} \times \mathbf{B}) \cdot d\mathbf{s} \\ &= \iint \left(\frac{\partial \mathbf{B}}{\partial t} - \text{rot}(\mathbf{v} \times \mathbf{B}) \right) \cdot d\mathbf{S} \end{aligned} \quad (6.3.10)$$

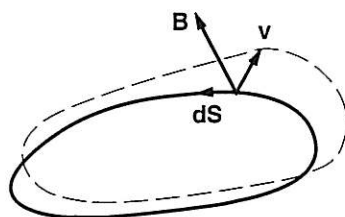


Figure 6.3.1. Illustration of the derivation of the law of flux conservation.

We then find that when the condition (6.3.8) is fulfilled so that the relation (6.3.9) holds, the right hand side of (6.3.10) vanishes identically, and the flux through the contour remains unchanged. This *flux conservation* also implies *field line conservation*, i.e. that volume elements of the medium which at one instant of time are located on a common magnetic field line, will be on a common magnetic field line at any other instant of time. This can be seen from the fact that any field line can be defined by the intersection between two surfaces which at all points are parallel to the local magnetic field vector. As the magnetic flux at any surface following the medium is unchanged, the two surfaces defining the field line will retain the property of being everywhere parallel to the magnetic field and hence always to define a magnetic field line.

6.4. The Critical Velocity

There occurs in the solar wind interaction with the coma of comets the phenomenon called the “critical velocity” or the *critical ionization velocity* (CIV). This phenomenon implies that a strong interaction takes place between the solar wind and the neutral gas in its way, in spite of the fact that the mean free path for collisions is so long that the interaction ought to be negligible.

The Hypothesis

The planets around the Sun are located at distance from the Sun, which form a famous series, known as *Titius–Bode’s law*: If one forms the number series 0, 3, 6, 12, 24, etc. and adds 4 to each of these numbers, a new number series, 4, 7, 10, 16, 28, etc. is obtained. It is found that this number series approximately represents the relative orbital radii of the planets.

As part of the theory of the origin of planets around the Sun and satellites around certain planets, Hannes Alfvén (1942) introduced the *hypothesis* that neutral gas falling in toward the early Sun was ionized and thereby stopped by the solar magnetic field, when the velocity of the fall reached a certain “critical velocity”, specific for each kind of particles. He found that the orbital radii of the planets as well as of the satellites could be explained rather well, if it was assumed that the critical velocity was that at which the kinetic energy of the neutral gas particles was equal to their ionization energy, *i.e.*

$$mv_{\text{crit}}^2/2 = eV_i \quad (6.4.1)$$

where m is the mass of the neutral gas particle. For hydrogen with the ionization energy 13.4 eV the critical velocity is $5.1 \cdot 10^4$ m/s. Heavy alkali metals have low critical velocity. For Ba, with an ionization energy of 5.2 eV and mass number 137, it is $2.7 \cdot 10^3$ m/s.

There was, however, no theoretical reason why the hypothetical strong interaction should take place when (6.4.1) was fulfilled. In the collision between an incident and a resident particle, it is the energy in the common center of gravity system of the two particles that is important. When (6.4.1) holds, the energy in the fixed system is equal to the ionization energy, but in the center of gravity system it is by necessity *less*. (If the relative velocity is v , and the particles are equally heavy, each particle has a velocity $v/2$, relative to the common center of gravity. In the center of gravity system, therefore, each has the energy $m(v/2)^2/2$, and the sum of their energies is $(1/2) \cdot mv^2/2$. If one of the particles is lighter, the energy in the center of gravity system becomes even less.) It should therefore be energetically impossible to achieve ionization. For this reason Alfvén’s cosmogonic theory was generally not taken seriously.

(*Cosmogony* is the science of the origin and evolution of the solar system, (cf. *cosmology* = the science of the origin and evolution of the Universe.) Alfvén

has emphasized that a theory of the origin of the planets must also be able to explain the origin of the satellite systems around certain planets, and has coined the word "*hetegony*" (from Greek *hetär* = companion).

Experimental Confirmation

Nevertheless, a few years later it was discovered, in an experiment at the Royal Institute of Technology, that the theoretically "impossible" interaction actually took place (Fahleson 1961). In this experiment a plasma was forced to rotate through a background of stationary neutral gas. With increasing power input the speed of rotation could be increased but only up to a certain limit, which could not then be exceeded until so much power was applied that the ionization became complete, and there was no neutral gas left to interact with. This velocity limit was found to be the one given in Section 6.4.1.

After this first experimental confirmation, the phenomenon was observed in a number of different experimental configurations, but remained unexplained. A Series of experiments have been performed at the Royal Institute of Technology, aimed at observing the phenomenon in as pure form as possible. Fig. 6.4.1 shows the result of such an experiment, where a plasma moving at supercritical velocity meets a neutral gas cloud, which is so thin that interaction should be negligible. The velocity is rapidly braked to a value equal to v_{crit} as soon as the plasma hits the neutral gas.

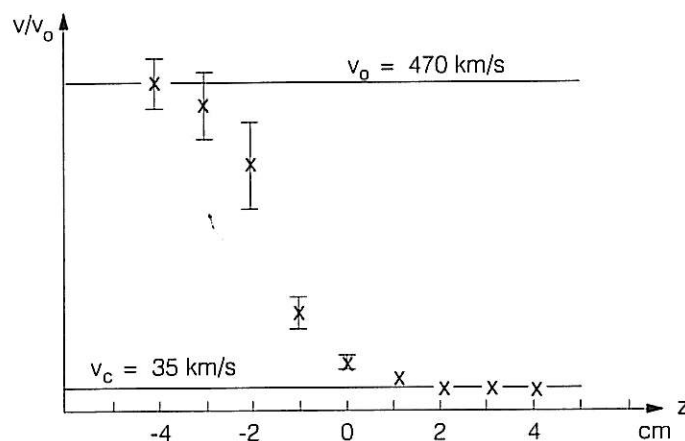


Figure 6.4.1. When the relative velocity between plasma and neutral gas exceeds a certain critical velocity, rapid braking and violent ionization takes place which cannot be explained by classical theory. The picture shows how a plasma moving with supercritical velocity encounters a neutral gas cloud and is rapidly braked to a velocity close to theoretical critical velocity (after Danielsson and Brenning 1975).

Explanation

On the basis of the experiments it has been possible to essentially clarify what happens (Raadu 1978): When the relative velocity between plasma and neutral gas reaches the critical velocity, an instability occurs which leads to strong electric field fluctuations, which transfers energy from the relative motion (between ions and neutral gas) to the electrons of the plasma. These can now in turn easily ionize the neutral gas particles. In order to trigger the instability it is sufficient that there exists the slightest ionization (*e.g.* as a result of cosmic radiation), which then grows like an avalanche. Certain problems remain, however, still to be solved (Brenning och Axnäs 1988).

Critical Velocity in Space Plasma

The fact that the critical velocity phenomenon occurs even in space plasma was discovered by an accident in the context of the Apollo program. When Apollo XIII had to return to Earth prematurely after an accident, the carrier rocket continued toward the Moon. When it hit the lunar surface with great force, dust and gas clouds were formed, which expanded out into the solar wind. Instruments left on the lunar surface by Apollo XII then showed a strongly enhanced flux of energetic ions. These were ions which were formed from the neutral gas cloud and subsequently accelerated in the electric field of the solar wind. The acceleration was expected, but what was surprising was that such a large part of the neutral gas cloud had become ionized. Analysis showed that the only explanation appeared to be the critical velocity phenomenon. (The solar wind velocity, 300 – 900 km/s, exceeds greatly the critical velocity for all ion species concerned.)

Interaction with Comets

It is of course to be expected that the critical velocity should also play a role when the solar wind encounters the comas of comets. The conclusion that this is the case has strong support from data from space probes, which have allowed observations of comets at close range (Galeev *et al.* 1986).

Interaction with the Interstellar Medium

Similarly critical velocity is considered to play a role for the interaction of the solar wind with the interstellar medium (Petelski *et al.* 1980), *i.e.* at the boundary of the heliosphere, the heliopause, but this is yet to be confirmed experimentally.

Active Experiments in Space Plasma

The phenomenon can nowadays also be studied in a controlled way by means of what are usually called *active experiments* in the vicinity of Earth, as is described more in detail in Chapter 10.

7. Stellar and Interstellar Plasma

Like the Sun, all stars consist of plasma. They also contribute plasma to interstellar space, partly by the outflux of plasma from certain stars and partly by short-wave electromagnetic radiation ionizing the surrounding interstellar medium.

Here follows a brief summary of some basic facts about the stars.

Brightness of Stars

The brightness of celestial objects as seen from the Earth is measured by a quantity m , called the *apparent magnitude*. Already classical astronomers like Hipparchus (190-125 B.C.) and Ptolemy (A.D. 100-178) used a classification system where the brightest stars were said to be of the first magnitude while the faintest stars visible to the naked eye were given the sixth magnitude. Due to the logarithmic sensitivity of the eye, the magnitude scale from one to six was roughly logarithmic. In modern terminology a difference of five magnitudes corresponds to a brightness ratio of one hundred which means that 1 magnitude is equivalent to a brightness ratio of $100^{1/5} \approx 2.512$. The apparent magnitude is then given by

$$M = -2.5 \log (\text{apparent brightness}) + \text{const.} \quad (7.1)$$

where the constant is chosen according to a certain convention. This gives Polaris (the Pole star) an apparent magnitude of $m = 2.0$ and Sirius, the brightest star in the sky, $m = -1.5$. If applied to the full Moon and the Sun this measure becomes -12.7 and -26.8, respectively. It is to be noticed that the fainter the object appears to be, the larger is the apparent magnitude. The faintest celestial objects observed with modern techniques have apparent magnitudes in excess of 29.

The apparent magnitude of a celestial object depends on both its intrinsic luminosity and its distance from the Earth. To get a measure of the intrinsic luminosity of an object one has defined the *absolute magnitude*, M , as the apparent magnitude the object would have if observed at a distance of 10 parsec (pc) (1 pc = 3.26 light years; definitions, see Appendix). One then finds that the Sun is of absolute magnitude $M = 4.8$, Sirius is $M = 1.0$, while Polaris is $M = -5.1$.

Above we have been speaking of *visual magnitudes*, the brightness as seen by the normal eye. The photographic plate has a different colour sensitivity and hence registers the celestial object differently. This gives the *photographic magnitudes*. By using filters one may also observe in different wavelength regimes, like infrared and ultraviolet, and each of these has its own magnitude system.

Spectral Classes

Stars with different surface temperature emit light with different spectral characteristics. On the basis of this, the stars have been grouped into spectral classes, denoted by the following letters.

O, B, A, F, G, K and M

and their subclasses denoted by attached numbers. Thus, spectral class O represents the hottest stars (surface temperature more than 20 000 K) and M the coolest (surface temperature about 3 000 K).

The Hertzsprung–Russell Diagram

If the stars are placed in a diagram with the surface temperature on the horizontal axis and light emission on the vertical axis, a diagram called the *Hertzsprung–Russell diagram* is obtained. An example is shown in Fig. 7.1.

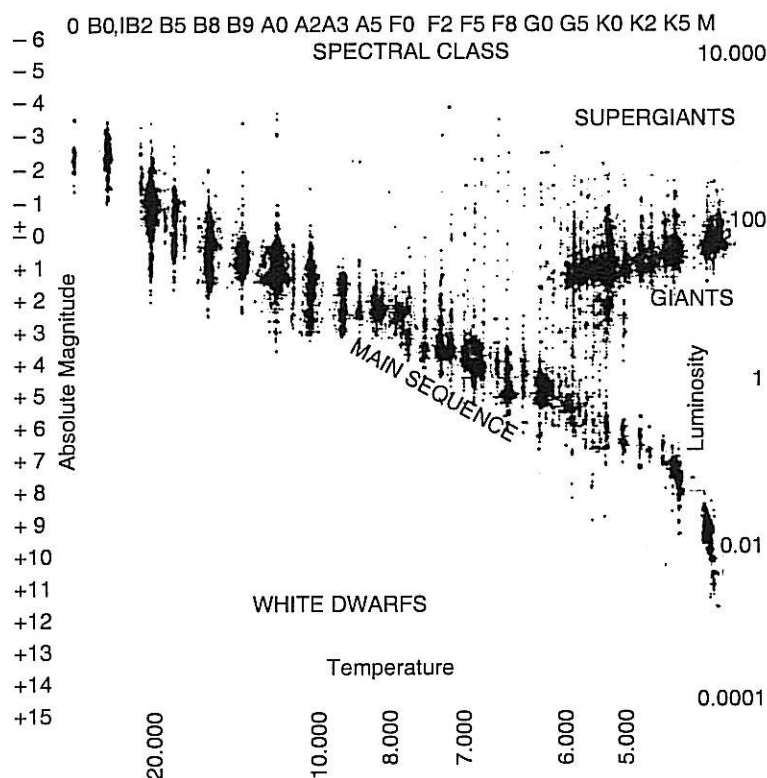


Figure 7.1. Hertzsprung–Russell diagram (after Wallenqvist 1958).

The position of stars in the Hertzsprung–Russell diagram can be used for classification, illustrated in Fig. 7.2.

The Sun

The Sun is located in the middle of the Hertzsprung–Russell diagram, in what is called the main sequence. During its evolution from birth to death the star describes a certain trajectory in the diagram.

Magnetic Stars

A particularly interesting category of stars is that of *magnetic* stars, which are characterized by their extremely strong magnetic fields (more than 3.4 T has been measured in the star HD 215441). They are usually found among rapidly rotating A-stars with a surface temperature of more than 10 000 K. The observed magnetic

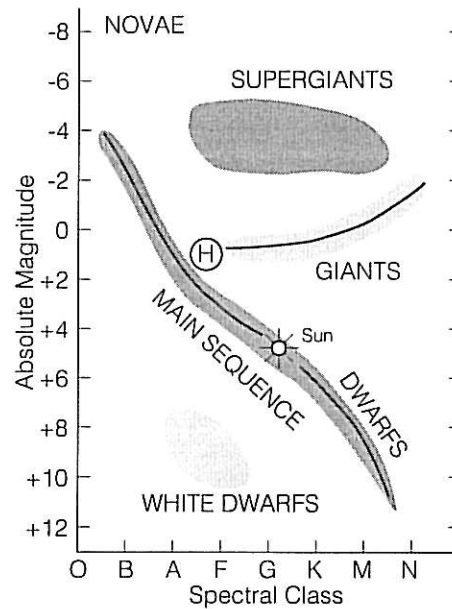


Figure 7.2. Classification of stars (after Wallenqvist 1958).

field strength usually shows large time variations (periodic or irregular) probably connected with the stars rotation.

Pulsars

The pulsars are objects which emit short (about 1 ms) pulses of radio emission, which recur strictly periodically with a recurrence frequency, which can vary between a few thousands of a second and a few seconds. In the cases where pulsars have been identified optically, it has been found that they coincide with the remnants of supernovas. (The most famous is the 30 ms pulsar in the Crab nebula, the result of the supernova in AD 1054.)

The pulsars are considered to be neutron stars, *i.e.* stars made up of nuclear matter (about 55 % neutrons and 45 % protons) with a mass of about 10^{57} neutron masses (about equal to the Sun's mass) and a radius of about 10 km, *i.e.* a density of more than 10^{14} g/cm³.

Stellar Wind Regions

Since a large part of the stars are hotter than the Sun, one must expect that they, too, emit plasma in the form of *stellar winds* and surround themselves with "astrospheres" corresponding to the Sun's heliosphere. These regions are thus populated with hot plasma directly from the star concerned.

Interstellar Matter

The stellar system, *The Milky Way*, of which our Sun is a part, is a *spiral galaxy*, whose diameter in the plane of the spiral is about 100 000 light years (10^{21} m) and transverse to it about 10 000 light years (10^{20} m). It contains about 10^{11} stars, the average distance between them being about 10 light years (10^{17} m).

The fact that there is also matter between the stars manifests itself in absorption lines in the stellar spectra, in some cases it also manifests itself through the reemission of light from strong stars (gas or reflection nebulae).

The occurrence of hydrogen gas can be mapped by using the 21-cm-line in the radio spectrum of hydrogen. Also molecular hydrogen exists, which manifests itself by emission of a spectral line ($\lambda = 2.6$ mm) which is excited in the CO-molecule in collisions with molecular hydrogen (Morris and Richard 1982).

In our own galaxy the total amount of atomic and molecular hydrogen are nearly equal (Sanders *et al.* 1984) but differently distributed. (Molecular hydrogen is dominant in the inner part of the galaxy, and concentrated to clouds with 1–30 solar masses, 20–80 parsec).

Interstellar matter, too, is in the plasma state, although often cool and weakly ionized (small relative number of charged particles). However, even when only a very small fraction of the particles is ionized, the macroscopic behaviour of the medium becomes in many important respects similar to that of the plasma. This is because the charged particles interact with long-range forces (the Coulomb force) and therefore usually have a much greater effective collision cross-section than the neutral gas particles (cf. Section 1.3).

Because of the great average distance between the stars, the above-mentioned stellar wind regions can only constitute a vanishingly small part of interstellar space. Much larger volumes are ionized as a consequence of the short-wave radiation of the stars. In interstellar space there are two types of interstellar plasma environments, called *HI*- and *HII*-regions respectively.

HII-regions are formed as a result of the radiation from a star in the UV and X-ray ranges. The ionization is nearly complete, the density about 1 cm^{-3} in the spiral arms of the galaxy, considerably less outside. The temperature is estimated at 10 000 K. Around stars of spectral classes O and B, *HII*-regions can extend to 50–500 light years.

HI-regions are found outside the reach of the stellar UV-radiation. The density is about the same as in the *HII*-regions, but the temperature can be as low as about 50 K, and as a consequence the dominating element, the hydrogen gas, is unionized, while certain other elements, *i.e.* Ca, are ionized, so that the gas still constitutes a plasma.

The interstellar plasma participates in the rotation of the galaxy (with a period of the order of $2 \cdot 10^8$ years) and in addition exhibits irregular (“turbulent”) motions with velocities of the order of magnitude of 10 km/s.

Interstellar Magnetic Fields

The interstellar plasma, too, is magnetized, which means that it carries electric currents. The existence of an interstellar magnetic field was suggested in 1937

by Alfvén to explain the isotropy of cosmic radiation, and has since been evoked in theories of the acceleration of the cosmic radiation. Early observations of the polarisation of star light was interpreted as a result of the interstellar magnetic field (magnetic dust grains become lined up by the magnetic field). An important method for mapping interstellar magnetic fields is analysis of the Faraday rotation (cf. Section 3.4) of the radio emission from pulsars. Out of this one has estimated that the large-scale magnetic field has a typical strength of $0.2 - 0.3$ nT and that the magnitude of the fluctuations is about 0.2 nT. Similar conditions have been found in other galaxies. A survey of the interstellar magnetic field has been given by Ruzmaikin *et al.* (1988)

8. Intergalactic Plasma

Space between the galaxies is called *intergalactic* space. The average distance between galaxies is estimated to be about 10 million light years (10^{23} m). The part of the universe that is until now observed is sometimes referred to as the *metagalaxy*, implying that other similar entities exist. The contribution of the stars to the average mass of the metagalactic space is estimated at less than one proton mass per cubic meter. There is, however, also diffuse intergalactic matter, among other things in the form of material bridges between neighbouring galaxies. Recently diffuse arches have been detected which surround galactic groups and have dimensions of more than 300 000 light years (Robinson 1987).

The discovery that there is intergalactic plasma and that it is magnetized was made by radio astronomical means. It was found that certain galaxies are surrounded by a *halo* of radio noise in the frequency range around 100 MHz. The surprising explanation of the radio noise was proposed by Alfvén and Herlofson (1950), namely that it is synchrotron radiation from superrelativistic electrons. In the weak magnetic field, about 0.3 nT, in which they move, the non-relativistic gyro frequency is only about 8 Hz. In this case, however, the electrons have extremely high energy, about 3 GeV. Therefore they have a relativistic mass about 6 000 times as large as their rest mass, and a gyro frequency of only about 0.0015 Hz. In spite of this they radiate in the 100 MHz range! This is because a relativistic electron emits most of its radiation at frequencies which exceed the gyro frequency by a factor $(m/m_0)^3$. This is because the radiation is emitted in a narrow cone around the instantaneous direction of motion of the electron. The higher the energy the smaller is the opening angle of the instantaneous emission cone. A fixed observer therefore sees the radiation as brief “flashes” each time he happens to be in this narrow cone around the instantaneous direction of the tangent to the electron orbit.

Since there are electrons outside the galaxies, there must also be positive particles, probably protons, otherwise unreasonably large electric fields would be created.

Based on the redshift of the spectra of distant galaxies, interpreted as a Doppler effect, one has found that the galaxies move away from us with a velocity proportional to the distance. The proportionality factor is the *Hubble parameter* (about 50 km/s per 1 million parsec (Schechter 1986)). Velocities of above 90 per cent of the velocity of light have been observed. The common interpretation of the redshift is that the visible Universe expands uniformly which explains the proportionality between distance and velocity.

It has been pointed out by Alfvén (see for example Alfvén 1981, Chapter IV:9), that one cannot be sure that metagalactic space consists of ordinary matter only. It is known from the physics of elementary particles that there is a symmetry between *particles* (the proton, the electron etc) and *antiparticles* (the antiproton, the positron etc). A corresponding symmetry should therefore exist between *matter* (consisting of particles) and *antimatter* (consisting of antiparticles). In a theory of the Universe as a whole it is therefore unsatisfactory to assume that ordinary matter, *koino matter* (from Greek *koinos* = common), should have an exceptional

place, as the case is in previous cosmological theories. A theory of the metagalaxy has therefore been proposed by Klein and Alfvén (for a summary see Alfvén 1965, and for a popular presentation Alfvén 1966). In this it is emphasized that the metagalaxy is not necessarily equal to the Universe as a whole, which may possibly be infinite both in time and space. In this theory the metagalaxy is assumed to consist of a dilute plasma of both koino matter and antimatter – an *ambiplasma* (from lat. *ambi* = double). This ambiplasma contracts under the influence of its own gravity, and the density increases. At a density of about $0.01 \text{ particles cm}^{-3}$ the pressure of the radiation from the annihilation becomes so large that it brakes the contraction and turns it into the expansion that is still going on. During some phase of the evolution of the metagalaxy, there have also, under the influence of electromagnetic forces, occurred separation processes, which have led to local concentrations of koino respectively antimatter (which finally become purely koino or antimatter because the minority species is destroyed by annihilation). At the boundaries between matter and antimatter there are formed hot boundary layers, partly analogous to the Leidenfrost-layer that protects a drop of water on a (sufficiently) hot plate. Therefore each individual star consists of *either* only matter (like our Sun) or only antimatter. Whether a certain star (or a certain galaxy) consists of matter or antimatter cannot be immediately determined, because the radiation that is emitted by atoms respectively antiatoms is precisely the same. Different suggestions have been made about how to determine whether the Universe contains antimatter (*e.g.* if one could observe spectral lines of protonium, *i.e.* atoms consisting of one proton and one antiproton), but at the present time the problem is still unsolved.

It has previously been thought that the Universe was rather homogenous on the large scale. Recently, however, it has been discovered that there are essential inhomogeneities (Burns 1986) on scales of tenths of megaparsec ($1 \text{ megaparsec} = 3.26 \cdot 10^6 \text{ light years} \approx 3 \cdot 10^{22} \text{ m}$). This is a difficulty (see for *e.g.* Maddox 1989, Lerner 1990) for the *Big-Bang*-theory (see *e.g.* Weinberg 1977) of the origin of the Universe, and a support for the opinion that also electromagnetic forces must be considered in cosmological theories (see *e.g.* Alfvén 1981).

9. Cosmic Radiation

The *cosmic radiation* consists of extremely high energy particles – energies above 10^{20} eV has been observed. (The lower limit of what is called cosmic radiation is somewhat arbitrary but is usually put at 10^8 eV.) Like the particle radiation in the Earth's radiation belts it constitutes a risk factor for both humans and electronic systems which are in space for an extended period of time.

Rigidity

Even at the high energies characterizing the cosmic radiation the particle orbits are influenced by the Earth's magnetic field. (Even at 10^9 eV energy, where the primary radiation has its maximum, the orbital radius of curvature in the inner part of the magnetosphere is of the order of 10 000 km, and the radiation can approach the Earth only from certain directions. This was used in the early space age to draw qualitative conclusions of the outer parts of the geomagnetic field.) In cosmic ray research one has therefore traditionally used a measure of energy, called *rigidity*, which represents the resistance of the particle to magnetic deflection.

If a particle with the charge $q = Ze$ and relativistic velocity moves transverse to a magnetic field with strength B the orbital radius of curvature becomes

$$R = p/eZB \quad (9.1)$$

where p is the relativistic momentum, which is related to the total energy E and the rest mass m_0 according to

$$E = (p^2c^2 + m_0^2c^4)^{\frac{1}{2}} \quad (9.2)$$

Since the orbital radius of curvature is proportional to p and inversely proportional to Z , p/Z , or any number proportional to this ratio, is a measure of the particle's capability to withstand magnetic deflection. One has chosen to define rigidity, P , as

$$P = pc/eZ \quad (9.3)$$

The relation between rigidity and kinetic energy per nucleon in the most important part of the cosmic ray energy spectrum is shown in Fig. 9.1.

Spectra

The energy spectrum of the cosmic radiation can be represented in various different ways. The *differential spectrum in a certain direction* is the number of particles per unit of time, energy, solid angle and surface area, which passes a surface perpendicular to this direction. This flux usually given in the unit $(\text{MeV/nucleon})^{-1}$

steradian⁻¹m⁻²s⁻¹. The *integrated* spectrum is correspondingly the number per unit of solid angle and time of particles above a certain energy and is measured in steradian⁻¹m⁻²s⁻¹.

For determining the extent of protection needed in the context of space activities it is usually sufficient to know the *omnidirectional spectrum*. This, too, can be given either as a differential spectrum or as an integrated spectrum and is obtained from the spectra mentioned above by integrating over solid angle.

For transforming spectra to radiation doses behind specified radiation shields there are computer programs which are published (Seltzer 1980).

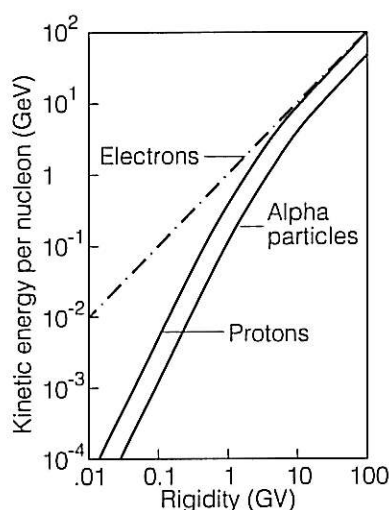


Figure 9.1. Relation between rigidity and kinetic energy per nucleon.
(After *Handbook of Geophysics and the Space Environment* 1985)

The Primary Cosmic Radiation

During its passage through the atmosphere (and through radiation shields) the cosmic radiation produces secondary particles because collisions with other particles lead to nuclear reactions. The cosmic radiation at the Earth's surface is therefore different than above the Earth's atmosphere. The latter is called the *primary* cosmic radiation.

The primary cosmic radiation at the Earth's orbit consists of 83 % protons, 13 % alpha particles, 1 % atomic nuclei with charge numbers greater than 2 and 3 % electrons.

Spectra of the two most important components in the primary cosmic radiation are shown in Fig. 9.2.

Solar Cycle Variations

The ability of the cosmic radiation to reach the orbit of the Earth is influenced by the state in interplanetary space, since this contains magnetic fields. At high solar activity the cosmic radiation is noticeably shielded and this gives rise to a solar cycle variation of the cosmic radiation which is illustrated in Fig. 9.3.

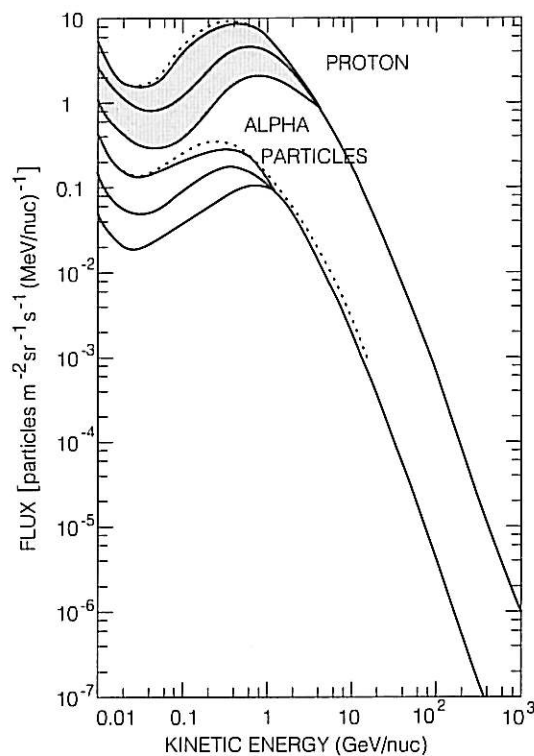


Figure 9.2. Differential energy spectra for protons and alpha particles. The proton spectrum in this figure has been multiplied by a factor of 5 in order to separate the curves (After *Handbook of Geophysics and the Space Environment* 1985.)

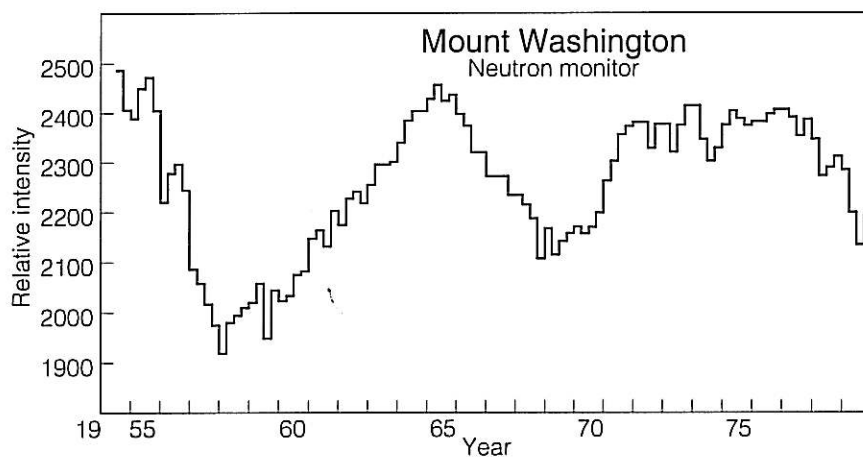


Figure 9.3. Solar cycle variation in the cosmic radiation as observed at the Earth's surface (After *Handbook of Geophysics and the Space Environment* 1985.)

Variations on Short Time Scales

Small variations also occur on short time scales. A 27-day periodic variation with an amplitude of a few per cent is caused by the fact that the state of the solar

wind in the vicinity of the position of the Earth is related to the solar rotation, which has a periodicity of 27 days.

At a given point on the surface of the Earth's daily variation is observed with an amplitude of some tenths to a few per cent, which depends on the position of the place relative to the average direction of the interplanetary magnetic field. (Also a small 12-hour variation, usually less than 0.1 % occurs.)

Transient variations occur as a result of solar flares. An important type is called "*Forbush decrease*" (after its discoverer, the American Scott Forbush). It can cause reductions in the cosmic ray intensity by a few to a couple of tenths of per cent (up to 35 % has been observed). This phenomenon is considered to be caused by rapidly moving magnetized solar wind plasma, temporarily sweeping away the cosmic radiation.

There are also increases as a result of particles emitted from the Sun (improperly called "solar cosmic rays") and as a result of shock waves acceleration of cosmic rays which are already present in interplanetary space.

Secondary Particles

The production of secondary particles begins at about 55 km altitude above the Earth's surface and reaches a maximum, called the *Pfotzer-maximum* at about 20 km. The secondary particles include protons, neutrons and mesons. (For example the integrated flux of μ -mesons has a maximum of about $400 \text{ m}^{-2} \text{ s}^{-1} \text{ steradian}^{-1}$ at an atmospheric depth of 200 g/cm^2 . Computer programs have been published for calculating the secondary radiation at a given place and at a given atmospheric depth (O'Brien 1970, 1979). Data on the secondary radiation has been compiled by Allkover and Grieder (1984).

The Neutron Albedo

Among the secondary particles of cosmic radiation neutrons are of particular interest for the following reason. Some of the neutrons that are formed in the atmosphere have such directions of motion that they are thrown out into space. The flux into space of neutrons created by the cosmic radiation is called the *neutron albedo*. (Albedo = equal to the ratio between the light diffusely reflected by an illuminated body and the incident light; the name derives from lat. *albus* = white.) Some of the albedo neutrons decay during their flight through the magnetosphere and the resulting protons become magnetically trapped. They constitute an important source (the first identified one) of the Earth's inner radiation belt.

Origin of the Cosmic Radiation

Except for protons from the Sun and electrons from Jupiter the origin of the cosmic radiation is still a matter of scientific debate. One of the best known theories

was proposed by Enrico Fermi. A magnetically trapped particle which is “squeezed” between two “magnetic mirrors” (cf. Section 4.8) increases its energy. This is called *Fermi acceleration* (cf. 4.8). The energy increase that can be achieved is, however, limited by the fact that the particle will sooner or later escape from the mirror. In order to reach the energies characteristic of cosmic radiation a multi-step process is needed. Fermi proposed that the cosmic ray particles are accelerated to their high energies through repeated collisions with “magnetic mirrors” which are “frozen-in” (cf. Chapter 6) in the interstellar plasma and moves with it. On reflexion against an approaching mirror the particle gains energy much like a ping-pong ball that is reflected by a racket. When reflected against a receding mirror the particle loses energy. It can be shown that the statistical net result of many random collisions will be successive increase of energy.

Whereas the cosmic radiation used to be generally considered to be of galactic origin, Hannes Alfvén has proposed that it may be largely produced within the heliosphere by a process called *magnetic pumping*. This is a combination of betatron acceleration (cf. 4.8) and elastic scattering. It takes place in the following way. Since the magnetic moment is an invariant, the kinetic energy of the particle motion transverse to the magnetic field will vary in direct proportion to the strength of this magnetic field according to (3.7.4). This process cannot provide any accumulated energy gain, because the energy gained is lost when the magnetic field strength returns to its old value. If, however, it is combined with elastic scattering, the following happens. When the magnetic field strength increases, part of the particles kinetic energy which corresponds to the motion transverse to the magnetic field, increases, but not that part of the energy that corresponds to the motion along the magnetic field. The elastic scattering, however, tends to distribute the kinetic energy equally between three degrees of freedom in the motion, and therefore transmits part of the “perpendicular” energy to “parallel” energy, which does *not* decrease when the field strength decreases. When the decrease of the magnetic field strength has taken place, energy is instead transferred to the “perpendicular” energy, which further increases when the magnetic field again increases. In this way it is possible, like in the case of stochastic Fermi acceleration, to increase the energy of the particles indefinitely. The “gedankenexperiment” (thought-experiment) just described is designed for a particularly simple kind of time variation, but it can be shown that a systematic increase of energy takes place even for an arbitrary time variation (see *e.g.* Alfvén and Fälthammar 1963, Section 2.7.4).

10. Space as a Laboratory

As mentioned already in the introductory chapter the vicinity of the Earth (the ionosphere and the magnetosphere) contain a rich variety of plasma populations with very different temperatures and densities. The density varies from more than 10^{12} m^{-3} (numerical density) at the maximum of the ionospheric F layer to less than 10^4 m^{-3} in the polar plumes of the magnetosphere, and (the equivalent) temperature from less than 1 000 K in the E layer of the magnetosphere to more than 10^9 K in the radiation belts of the Earth. A consequence of this variety is a profusion of natural plasma physical processes take place and in addition conditions are favourable for “active experiments”, where artificial particle or wave populations are injected and the response of the space plasma is studied.

Since the natural physical conditions in space have now been largely explored (“the exploration phase”) interest has grown in gaining a deeper understanding of the physics of the space plasma by means of active experiments (“the physics phase”).

Most active experiments performed so far can be divided into four main groups:

1. Injection of artificial clouds of plasma or neutral gas
2. Injection of particle beams of electrons or ions
3. Study of the effect of spacecraft on the ambient plasma
4. Injection of waves.

A few examples of such experiments will be given below. An up-to-date collection of papers on such experiments has been edited by Haerendel and Mendillo (1988).

Injection of Artificial Gas or Ion Clouds

The earliest use of artificial ion clouds was measurements of electric fields in the ionosphere. By detonating the mixture of an explosive and barium or strontium (both of which have low ionization energy) in a sunlit part of the ionosphere a gas cloud was formed which became ionized by the solar radiation. The freshly formed barium or strontium ions immediately started drifting under the influence of the electric field according to the formula (3.7.10). Since the magnetic field in the ionosphere is well known, one can use observations of the motion of the ion clouds to directly calculate the electric field. The determination of the motion of the ion cloud is further facilitated by the fact that one can at the same time observe the still unionized atoms of the cloud, the motion of which is not influenced by either the magnetic or the electric field. This was the first method of measuring electric fields in the ionosphere, but it has been used to a rather limited extent, since it can only be used in regions with special light conditions, namely near the terminator, and does not allow high time and space resolution.

In another early application of artificial clouds the shaped charge technique was used to inject barium clouds with very high velocity (10 km/s) *along* the geomagnetic field. With this method a very important discovery was made in 1976 (Haerendel 1976), which is illustrated in Fig. 10.1. It was found that about 700 seconds after the release at an altitude of about 7 000 km, the ion cloud suddenly was subjected to a sudden velocity change corresponding to an energy increase of about 7 keV per particle. The only explanation of this is that the cloud passes through a region with an electric field along the magnetic field lines. This provided undeniable proof that such fields exist in the space plasma. Such fields, the existence of which were long doubted, are considered to play an important role in the space plasma, among other things for the acceleration of the auroral primary particles (cf. Section 4.5).

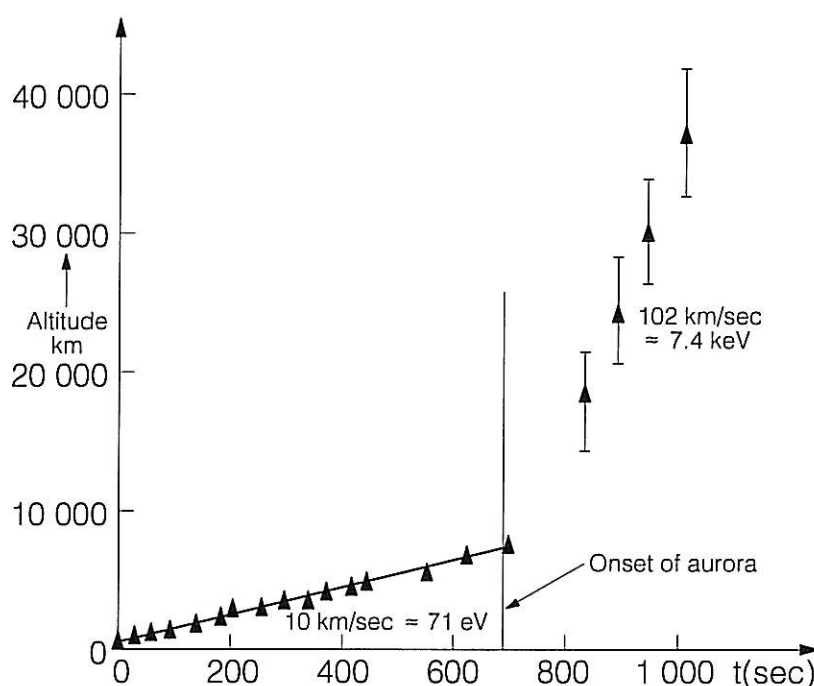


Figure 10.1. Motion of Ba^+ clouds along the geomagnetic field (Haerendel *et al.* 1976).

An ambitious international project for active experiments on magnetospheric scale was performed in 1984/85 – the AMPTE-project (Active Magnetospheric Particle Tracer Experiment), which included three different satellites and several experiments.

One of these experiments consisted of injection of lithium ions in front of the magnetopause. The idea was that these ions, which do not occur naturally in space in any appreciable quantity, might be detected inside the magnetosphere and reveal something about the mechanism for plasma penetration from the solar wind into the magnetosphere. It was, however, not possible to detect any lithium ions within the magnetosphere.

Another experiment consisted of injection of a number of barium clouds (each with about 10^{25} barium atoms) in the solar wind to simulate what happens in the diffuse gas envelope of the comet (koma, from Greek *kome* = hair). The explosion formed, as expected, a magnetic cavity by forcing aside the surrounding solar wind

plasma with its frozen-in magnetic field. But several unexpected phenomena were observed. Perhaps the most interesting was that the surrounding magnetic field returned to the cavity in a very short time, a few seconds to a few minutes, which was many times faster than expected according to theoretical predictions. This is one example among many which illustrates how observations and experiments in space have shown that the behaviour of matter in the plasma state can substantially deviate from what classical theories prescribe.

By means of suitable gases that cause strong electron absorption temporary “holes” have been punched in the ionosphere (even the exhaust gases from the rocket motors of the space shuttle can be used for this purpose.) This has been used both to make possible radio observations in frequencies that do not normally penetrate through the ionosphere, and to study the reestablishment of the ionospheric layers.

A particularly interesting category of experiments uses the ionospheric plasma for study of “the critical-velocity” phenomenon (cf. Section 6.4). This kind of experiments which are performed with a participation of scientists at the Royal Institute of Technology, will be described in a separate section later.

Injection of Particle Beams

Most experiments with particle beams have been used to study the plasma instabilities which such beams give rise to by their motion through the background plasma. They have also been used to produce artificial aurora and to look for electric potential barriers (for example electric double layers, cf. Section 4.5) in space.

Electron beams have been used on the European satellites GEOS 1 and GEOS 2 to measure electric fields by emitting an electron beam, which is received by detectors on the same satellite. The time, place and angle of incidence of the returning beam contains information on the electric field in the surrounding of the satellite. A German research group at the Max Planck Institut für Extraterrestrische Physik in Munich will use a further sophisticated version of this experiment on the next Swedish scientific satellite, Freja (see Chapter 11), and on the satellites in the ESA-project Cluster.

Attempts to develop an ion beam technique for measuring electric fields from sounding rocket in the ionosphere are being made in Japan.

Influence of Spacecraft on Ambient Plasma

Several studies have been made of the plasma wake behind rockets and even larger spacecraft, such as the space shuttle. One such experiment used small subsidiary satellite “in orbit” around the space shuttle.

A more ambitious project of this kind is the Tether-project, which uses a “tethered” satellite, which will be connected with the space shuttle with a cable of a length of up to one hundred kilometers. Two kinds of experiments will be made. In one of them the subsidiary satellite is hanging under the space shuttle in a cable 20 km long, whereby it can reach down to levels where the atmospheric friction is

too large for free-flying satellites. In the other it will be hanging *upwards* from the space shuttle in a cable of 100 km length. In this case the orbital motion across the geomagnetic field will induce an emf of many kilovolts between the ends of the cable and drive currents along the geomagnetic field lines to and from the underlying ionosphere. One can in principle also force a current in the opposite direction, and thereby add kinetic energy to the space shuttle.

Injection of Waves

Injection of radio waves of high intensity from transmitters on the ground has been used for local heating of the ionosphere and study of subsequent behaviour.

Microwaves emitted from spacecraft have been used to trigger various plasma instabilities in the ambient plasma.

Radio waves from satellites can also be used both for local density measurement and for remote sensing of plasma discontinuities, such as the plasmopause.

The CRIT-Experiments – an Example of Plasma Experiment in Space

One of the most intricate problems in modern plasma physics has been the phenomenon “critical velocity” in the interaction between plasma and neutral gas. The problem and its background is described in Section 6.4. The fact that the phenomenon occurs also in the near-Earth space plasma was demonstrated for the first time by Haerendel (1982), but it has also failed to occur in other similar experiments.

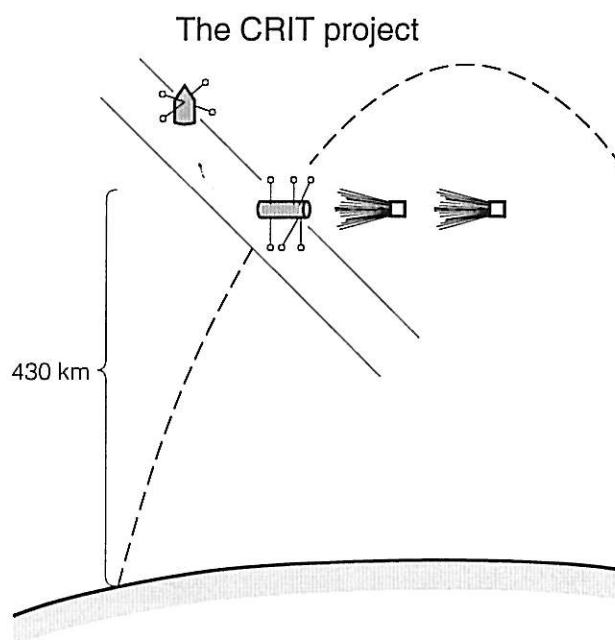


Figure 10.2. The CRIT-experiment

The CRIT-project, a collaborative project between scientific groups at University of Alabama, University of Alaska, Cornell University, Max Planck Institute and the Royal Institute of Technology, is the most comprehensive experiment so far. The experiment, which is illustrated in Fig 10.2, is performed in such a way that a rocket with double instrument payloads ("mother-daughter") is launched to an altitude of about 430 km from NASA's rocket range on Wallops Island in Virginia. During the upleg of the flight the smaller instrumented payload the "daughter" is separated in such a way that at the same time as it moves away, it always stays on the same magnetic field line as the main payload. During the upleg of the flight the main payload is turned by means of attitude controlling gas jets in such a way that its axis form a suitable angle to the magnetic field. After that two cannisters of barium-doped shaped charges are injected. When they have reached suitable distances (a few kilometers) from the instrumented payload, the charges are detonated, first one and then the other, in the direction of the instrumented payload, which will then be located in the middle of the region where the artificial barium clouds interact with the surrounding ionosphere. At the same time as the main payload measures electric and magnetic fields, waves and particles in the interaction region itself, the daughter payload checks what happens further away along the same magnetic field line, for example passage of energetic particles or Alfvén waves from the region of interaction.

The first experiment in this project was performed on the 13th of May 1986 and revealed the occurrence of still unexplained very strong (700 mV/m) electric fields along the geomagnetic field lines. In order to explain these, numerical simulations with a supercomputer are being performed at the computer center in Linköping.

A second experiment was performed on the 4th of May 1989 and has given equally interesting results which are still being analyzed at the institutions involved.

Experiments in the space plasma have often revealed that conclusions built on purely theoretical calculations can be dramatically misleading. One conclusion of this is that it is only by first understanding the space plasma in the available parts of space that one can hope to correctly interpret astrophysical observations which take place in regions that are far beyond the reach of direct measurements.

Bibliography

Handbooks etc.

- Allen, C.W., *Astrophysical Quantities*, 3rd Ed., Athlone Press, London, 1973
- Bruzek, A. and Durrant, C. J., Editors, *Illustrated Glossary for Solar and Solar-Terrestrial Physics*, D. Reidel Publishing Co., Dordrecht, Holland, 1977.
- Daly, E. J., The Evaluation of Space Radiation Environments for ESA Projects, *ESA Journal*, 12, 229-247, 1988.
- Folkestad, K., Lectures for EISCAT Personnel, Vol. I, EISCAT Technical Note 79/19, Kiruna, Sweden 1979.
- Folkestad, K., Lectures for EISCAT Personnel, Vol. II, EISCAT Technical Note 81/29, Kiruna, Sweden 1981.
- Johnson, F. S., Editor, *Satellite Environment Handbook*, Stanford Univ. Press, 1965.
- Jursa, A. S., Sci. Ed. *Handbook of Geophysics and the Space Environment*, Air Force Geophysics Laboratory, 1985 (Distributed by National Technical Information Service, 5285 Port Royal Road, Springfield, Virginia, USA.)
- Seltzer, S., SHIELDOSE: A Computer Code for Space-shielding Radiation Dose Calculations, NBS Technical Note 1116, National Bureau of Standards, U S Department of Commerce, Washington, D C, May 1980.
- Smith, R. E., and West, G. S., Space and Planetary Environment Criteria. Guidelines for Use in Space Vehicle Development. NASA Technical Memorandum 82478, 1973.
- Tandberg-Hanssen, E., Prominences, pp. 97-109, in *Illustrated Glossary for Solar and Solar-Terrestrial Physics*, Eds. Bruzek, A., and Durrant, C. J., D. Reidel Publ. Co., Dordrecht, Holland, 1977.
- U S Standard Atmosphere 1976*, NOAA, NASA, USAF, Washington D.C., 1976.

Monographs

- Athay, R. G., *The Solar Chromosphere and Corona: Quiet Sun*, Reidel Publ. Co., Dordrecht 1976.
- Akasofu, S.-I. and Kamide, A. Y., *The Solar Wind and the Earth*, Terra/Reidel Publ. Co., 1987.
- Akasofu, S.-I. and Chapman, S., *Solar-Terrestrial Physics*, Oxford University Press 1972.

- Alfvén, H. and Arrhenius, G., *Structure and Evolution of the Solar System*, NASA SP-345, Washington, D. C., 1976.
- Alfvén, H., *Cosmic Plasma*, D. Reidel Publ. Co., Dordrecht, Holland, 1981.
- Alfvén, H., *Cosmical Electrodynamics*, Oxford University Press, 1950.
- Alfvén, H., and Fälthammar, C.-G., *Cosmical Electrodynamics, Fundamental Principles*, Oxford University Press, 1963.
- Bolin, B., Huvuddragen av den övre atmosfärens fysik, MISU (Meteorologiska Institutionen vid Stockholms Universitet), Stockholm, 1969.
- Cravens, Thomas, E., *Physics of Solar System Plasmas*, Cambridge University Press, 1997.
- Egeland, A., Ø. Holter and A. Omholt, *Cosmical Geophysics*, Universitetsforlaget, Oslo, 1973.
- Folkestad, K., Lectures for EISCAT Personnel, Part I and Part II, EISCAT Technical Notes 79/19, 1979 and 81/29, 1981.
- Friedman, H., *Sun and Earth*, Scientific American Library, New York, 1986.
- Gombosi, Tomas, E. Gombosi, *Physics of the Space Environment*, Cambridge University Press, 1998.
- Gibson, E. G., *The Quiet Sun*, NASA SP-303, 1973.
- Hargreaves, J. K., *The Upper Atmosphere and Solar-Terrestrial Relations*, Van Nostrand Reinhold Co., 1979.
- Helliwell, R. A., *Whistlers and Related Ionospheric Phenomena*, Stanford University Press, 1965.
- Hultqvist, B., *Introduktion till Geokosmofysiken*, Natur och Kultur, 1967.
- Jensen, E. and Engvold, O., *Solen - en införelse i moderne solfysikk*, Universitetsforlaget, Oslo 1977.
- Kallenrode, May-Britt, *Space Physics*, Springer, 1998.
- Lyons, L. R., and Williams, D. J., *Quantitative Aspects of Magnetospheric Physics*, D. Reidel Publ. Co., Dordrecht, Holland, 1984.
- Menzel, D. H., *Our Sun*, Harvard University Press, 1959.
- Northrop, T. G., *The Adiabatic Motion of Charged Particles*, Interscience 1963.
- Panofsky, W. K. H. and Phillips, M., *Classical Electricity and Magnetism*, Addison-Wesley Publ. Co., London 1962.
- Ruzmaikin, A. A., Shukurov, A. M., and Sokoloff, D. D.: *Magnetic Fields of Galaxies*, Kluwer Academic Publishers, Dordrecht, Holland, 1988.
- Störmer, C.: *The Polar Aurora*, Clarendon Press, Oxford, 1955.
- Tandberg-Hanssen, E., and Emslie, A. G., *The Physics of Solar Flares*, Cambridge Astrophysics Series, 1988.

Tascione, T. F., *Introduction to the Space Environment*, Orbit Book Co., Malabar, Florida, 1988.

Wallenqvist, A., *Astronomi*, Bonniers, Stockholm 1958.

Popular Publications

Akasofu, S.-I., The Dynamic Aurora, *Sci. Amer.*, *260*, No. 5, pp. 54-63, May 1989.

Alfvén, H. The Plasma Universe, *Phys. Today* *39*, No. 9, pp. 22-27, 1986.

Alfvén, H., *Världen - Spegelvärlden, Kosmologi och Antimateria*, Aldus-Bonniers, 1966.

Balick, B., The Shaping of Planetary Nebulae, *Sky and Telescope*, *73*, No. 2, pp. 125-130, February 1987.

Balsiger, H., Fechtig, H., and Geiss, J., A Close Look at Halley's Comet, *Sci. Amer.*, *259*, No 6, pp. 62-69, December 1988.

Belcher, J. W., Bridge, H. S., Bagnal, F., Coppi, B., Divers, O., Eviatar, A., Gordon, G. Jr., Lazarus, A. J., McNutt, R. L. Jr., Plasma Observations Near Neptune: Initial Results from Voyager 2, *et al.*, *Science*, *246*, pp. 1478-1482, 1989.

Burns, J. O., Very Large Structures in the Universe, *Sci. Amer.*, *255*, No 1, pp. 30-39, July 1986.

Eather, R. M., *Majestic Lights*, Amer. Geophys. Union, Washington, D.C., 1980.

Fredga, K., A Viking Explores the Secrets of the Northern Lights, *The SAAB-Scania Griffin* (A Technical Journal Published Annually by the SAAB-Scania Group), pp. 2-9, 1987.

Friedman, H., *Sun and Earth*, Scientific American Library, New York, 1986.

Fälthammar, C.-G., Les aurores polaires, *La Recherche*, *3*, pp. 537-545, 1972.

Fälthammar, C.-G., Den joniserade materien - plasmat i kosmos och i laboratoriet, *Documenta Nr 30*, The Royal Academy of Sciences, Stockholm 1978.

Helfant, D., Bang: The Supernova of 1987, *Physics Today*, *40*, No. 8, pp. 25-32, 1987.

Hones, E. W., The Earth's Magnetotail, *Sci. Amer.*, *254*, No 3, pp. 32-39, March 1986.

Ingersoll, A. P., Uranus, *Sci. Amer.*, *256*, No 1, pp. 30-37, January 1987.

Kinoshita, J., Neptune, *Scientific American*, *261*, No. 5, pp. 60-69, November 1989.

Krimigis, S. M. *et al.*, The Magnetosphere of Uranus: Hot Plasma and Radiation Environment, *Science*, *233*, pp. 97-102, 1986.

Krimigis, S. M., Armstrong, T.P., Axford, I. A., Bostrom, C. O., Cheng, A. F. and Van Allen, J. A., Hot Plasma and Energetic Particles in Neptunes Magnetosphere, *Science*, *246*, pp. 1483-1489, 1989.

- Lanzerotti, L. J. and Krimigis, S. M., Comparative Magnetospheres, *Physics Today*, 38, No.11, pp. 24-34, November 1985.
- Lanzerotti, L. J. and Uberoi, C., The Planets' Magnetic Environments, *Sky and Telescope*, 77, No.2, pp. 149-152, February 1989.
- Lanzerotti, L. J., Earth's Magnetic Environment, *Sky and Telescope*, 76, No. 4, pp. 360-362, October. 1988.
- Leibacher, J. W. *et al.*, Helioseismology, *Sci. Am.*, 253, No 3, pp. 34-43, September 1985.
- Lerner, E., *The Big Bang Never Happened*, Random House 1990.
- MacKeown, P. and Weekes, T. C., Cosmic Rays from Cygnus X-3, *Sci. Amer.*, 253, pp. 40-49, 1985.
- Ness, N. F., Acuna, M. H., Burlaga, L. F., Connerney, J. E. P., Lepping, R. P. and Neubauer, F. M., Magnetic Fields at Neptune, *Science*, 246, pp. 1473-1478, 1989.
- Paresce, F., and Bowyer, S., The Sun and the Interstellar Medium, *Sci. Amer.*, 255, No 3, pp. 89-94, September 1986.
- Parker, E. N., The Sun, *Scientific American*, 233, pp. 42-57, 1975.
- Parker, E. N., Why Do Stars Emit X-rays?, *Physics Today*, 40, pp. 36-42, 1987.
- Robinson, L. J., Giant Galactic Arcs, *Sky and telescope*, 73, No. 4, p. 379, April 1987.
- Robinson, L. J., Halley Finale, *Sky and Telescope*, 72, No. 2, pp. 118-123, August 1986.
- Robinson, L. J., The Sunspot Cycle: The Tip of the Iceberg, *Sky and Telescope*, 73, No. 6, pp. 589-591, June 1987.
- Sagdeev, R. Z., and Galeev, A. A., Comet Halley and the Solar Wind, *Sky and Telescope*, 73, No. 3, pp. 252-255, March 1987.
- Schechter, B., Measuring the Hubble Constant, *Physics Today*, January. 1986, pp. 20-21.
- Shaham, J., The Oldest Pulsars in the Universe, *Sci. Amer.*, 256, No 2, pp. 34-40, February 1987.
- Soderblom, L. A. *et al.*, The Jupiter System Through the Eyes of Voyager 1, *Science*, 204, pp. 951-971, 1979.
- Van Allen, J. A., *Origins of Magnetospheric Physics*, Smithsonian Institution Press, 1983.
- Weinberg, S., *The First Three Minutes - A Modern View of the Origin of the Universe*, Basic Books Inc., New York 1977.
- Weissman, P. R., Realm of the Comets, *Sky and Telescope*, 73, No. 3, pp. 238-241, March 1987.
- West, R. M., Europe's Astronomy Machine, *Sky and Telescope*, 75, No. 5, pp. 471-481, May 1988.

White, N. E., New-Wave Pulsars, *Sky and Telescope*, 73, No. 1, pp. 22-24, January 1987.

Wyller, A., LEST (Large Earth-Based Solar Telescope) - An Overview, Institut for Teoretisk Astrofysikk, Universitetet in Oslo, 1986.

Scientific Papers

Acuna, M. H., *et al.*, Physics of the Jovian and Saturnian Magnetospheres, *Space Sci. Rev.*, 35, pp. 269-292, 1983.

Akasofu, S.-I., The Development of the Auroral Substorm, *Planet. Space Sci.* 12, pp. 273-282, 1964.

Albert, R. D, Nearly Monoenergetic Electron Fluxes Detected During a Visible Aurora, *Phys. Rev. Letters*, 18, pp. 369-372, 1967.

Alfvén, H., A Theory of Magnetic Storms and the Aurorae, *Kungl. Svenska Vetenskapsakademiens Handlingar*, 18, No. 3, p.1, 1939.

Alfvén, H., A Theory of Magnetic Storms and the Aurorae, II, The Aurorae, *Kungl. Svenska Vetenskapsakademiens Handlingar*, 18, No. 9, pp 1, 1939.

Alfvén, H., A Theory of Magnetic Storms and the Aurorae, III, The Magnetic Disturbances, *Kungl. Svenska Vetenskapsakademiens Handlingar*, 18, No. 9, p. 18, 1939.

Alfvén, H., Existence of Electromagneto-Hydrodynamic Waves, *Nature*, 150, p. 405, 1942.

Alfvén, H., On the Cosmogony of the Solar System, *Stockholms Observatoriums Annaler*, I, 14, No. 2, 1942.

Alfvén, H., and Herlofson, N., Cosmic Radiation and Radio Stars, *Phys. Rev.*, 78, pp. 616, 1950.

Alfvén, H., On the Theory of Magnetic Storms and Aurorae, *Tellus*, 10, pp. 104-116, 1958.

Alfvén, H., On the Theory of Comet Tails, *Tellus*, IX, pp. 92-96, 1959.

Alfvén, H., Antimatter and the Development of the Metagalaxy, *Rev. Mod. Phys.*, 37, pp. 652-665, 1965.

Allkover, O. C., and Grieder, P. K. F., Cosmic Rays on Earth, *Physics Data 25-1*, Fachinformationszentrum Energie Physik, Matematik GmbH, Karlsruhe, 1984.

Axford, W. I., Some Aspects of the Structure and Dynamics of the Terrestrial Magnetosphere, p. 5 in *Natural Phenomena below 30 kc/s*, Ed. D. F. Bleil, Plenum Press, 1964.

Becker, J. M., Quiet Photosphere and Chromosphere, pp. 21-34 in *Illustrated Glossary for Solar and Terrestrial Physics*, Red. A. Bruzek and C. J. Durrant, Reidel Publ. Co., Dordrecht 1977.

- Biermann, L., Kometenschweife und Solare Korpuskularstrahlung, *Z. Astrophys.* **29**, 274-286, 1951.
- Biermann, L., and Lüst, R., Comets: Structure and Dynamics of Tails, pp. 618-638 in *The Moon, Meteorites and Comets*, University of Chicago Press, 1963.
- Block, L. P., Fälthammar, C.-G., Lindqvist, P.-A., Marklund, G., Mozer, F. S., Pedersen, A., Potemra, T. A. and Zanetti, L. J., Electric Field Measurements on Viking: First Results, *Geophys. Res. Lett.*, **14**, pp. 435-438, 1987.
- Boström, R., Koskinen, H., and Holback, B., Low Frequency Waves and Solitary Structures Observed by Viking, *ESA-SP-275*, pp. 185-192, 1987.
- Brenning, N., and Axnäs, I., Critical Ionization Velocity Interaction: Some Unsolved Problems, *Astrophys. Space Sci.*, **144**, pp. 15-30, 1988.
- Bruzek, A., Spots and Faculae, pp. 71-79, in *Illustrated Glossary for Solar and Solar-Terrestrial Physics*, Eds. Bruzek, A., and Durrant, C. J., D. Reidel Publ. Co., Dordrecht, Holland, 1977.
- Burton, R. K., McPherron, R. C., and Russell, C. T., The Terrestrial Magnetosphere: A Half Wave Rectifier of the Interplanetary Electric Field, *Science*, **189**, pp. 717-718, 1975.
- Burton, W. B., The Morphology of Hydrogen and Other Tracers in the Galaxy, *Ann. Rev. Astron. Astrophys.*, **14**, pp. 275-306, 1976.
- Cahill, L. J. and Amazeen, P. E., The Boundary of the Geomagnetic Field, *J Geophys Res.*, **68**, pp. 1835-1843, 1963.
- Carlqvist, P., and Alfvén, H., Energy Source for the Solar Wind, *Astrophys. Space Sci.* **71**, pp. 203-209, 1980.
- Carlqvist, P., Current Limitation and Solar Flares, *Solar Physics* **7**, pp. 377-392, 1969.
- Carlqvist, P., On the Acceleration of Energetic Cosmic Particles by Electrostatic Double layers, *IEEE Transactions on Plasma Science*, *PS-14*, p. 794, 1986.
- Carpenter, D. L., Whistler Evidence of a "Knee" in the Magnetospheric Ionization Density Profile, *J. Geophys. Res.*, **68**, pp. 1675-1682, 1963.
- Chiu, Y. T., and Cornwall, J. M., Electrostatic Model of a Quiet Auroral Arc, *J. Geophys. Res.*, **85**, pp. 543-556, 1980.
- Cloutier, P. A., Daniell, R. E., Dessler, A. J., and Hill, T. W., A Cometary Model for Io, *Astrophys. Space Sci.*, **55**, pp. 93-112, 1978.
- Danielsson, L. and Brenning, N., Experiments on the Interaction Between a Plasma and a Neutral Gas, *Phys. Fluids*, **18**, pp. 661-671, 1975.
- Dessler, A. J., Solar Wind and Interplanetary Magnetic Field, *Rev. Geophys.* **5**, pp. 1-41, 1967.
- Evans, D. S., The Observation of a Near-Monoenergetic Flux of Auroral Electrons, *J. Geophys. Res.*, **73**, No. 7, pp. 2315-2323, 1968.
- Fahleson, U., Experiments with Plasma Moving Through Neutral Gas, *Phys. Fluids*, **4**, No. 1, pp. 123-127, 1961.

Frank, L. A., and Craven, J. D., Imaging Results from Dynamics Explorer 1, *Rev. Geophys.*, *26*, pp. 249-283, 1988.

Frohlich, C., Contemporary Measures of the Solar Constant, in *The Solar Output and its Variation*, Ed. O. R. White. Colorado Associated University Press, Boulder 1977.

Fälthammar, C.-G., The Solar Wind, Chapter 7 in *Cosmical Geophysics*, Ed. A. Egeland, Ø. Holter and A. Omholt, Universitetsforlaget, Oslo, 1973a.

Fälthammar, C.-G., Motion of Charged Particles in the Magnetosphere, Chapter 9 in *Cosmical Geophysics*, Ed. A. Egeland, Ø. Holter and A. Omholt, Universitetsforlaget, Oslo, 1973b.

Fälthammar, C.-G., Generation Mechanisms for Magnetic-Field Aligned Electric Fields in the Magnetosphere, *J. Geomagn. Geoelectr.*, *30*, pp. 419-434, 1978.

Fälthammar, C.-G., Magnetic-Field Aligned Electric Fields, *ESA Journal*, *7*, pp. 385-400, 1983.

Fälthammar, C.-G., Magnetosphere-Ionosphere Coupling, *ESA Special Report ESA-SP 235*, pp. 107-133, 1985.

Fälthammar, C.-G., Magnetosphere-Ionosphere Interactions - Near-Earth Manifestations of the Plasma Universe, *Astrophysics and Space Science*, *144*, pp. 105-133, 1988a.

Fälthammar, C.-G., Astrophysical Significance of Observations and Experiments in the Earth's Magnetosphere, *ESA Special Report ESA-SP 285*, Vol. 1, pp. 9-19, 1988b.

Fälthammar, C.-G., Electric Fields in the Magnetosphere - A Review, *Planet. Space Sci.*, *37*, pp. 899-914, 1989a.

Fälthammar, C.-G., Electrodynamics of Cosmical Plasmas - Some Basic Aspects of Cosmological Importance, *IEEE Transactions on Plasma Science*, *17*, pp. 174-185, 1989b.

Fälthammar, C.-G., Akasofu, S.-I., and Alfvén, H.: The Significance of Magnetospheric Research for Progress in Astrophysics, *Nature*, *275*, pp. 185-188, 1978.

Fälthammar, C.-G., Block, L. P., Lindqvist, P.-A., Marklund, G., Pedersen, A., Mozer, F. S., Preliminary Results from the D.C. Electric Field Experiment on Viking, *Ann. Geophys.*, *5A*, pp. 171-176, 1987.

Galeev, A. A., Gringauz, K. I., Klimov, S. I., Remisov, A. P., Sagdeev, R. Z., Savin, S. P., Sokolov, A. Yu., Verigin, M. I., and Szego, K., Critical Ionization Velocity effects in the Coma of Comet Halley: Measurements by Vega-2, *Geophys. Res. Lett.*, *13*, pp. 845-848, 1986.

Gold, T., Motions in the Magnetosphere of the Earth, *J. Geophys. Res.*, *64*, pp. 1219-1224, 1959.

Haerendel, G., Rieger, R., Valenzuela, A., Föppl, H., Stenbaek-Nielsen, H. C., and Wescott, E. M., First Observations of Electrostatic Acceleration of Barium Ions Into the Magnetosphere, *ESA-SP 115*, pp. 203-211, 1976.

- Hale, G. E., On the Probable Existence of a Magnetic Field in Sunspots, *Astrophys J.*, 28, p. 319, 1908.
- Hale, G. E., Seares, F. H., von Maanen, A., and Ellerman, F., The General Magnetic Field of the Sun, *Astrophys. J.*, 47, p. 206, 1918.
- Hale, L. C., Croskey, C. L. and Mitchell, J. D., Measurements of Middle-Atmosphere Electric Fields and Associated Electrical Conductivities. *Geophys. Res. Lett.* 8, pp. 927-930, 1981.
- Hanson, W. B., Structure of the Ionosphere, p. 20 in *Satellite Environment Handbook*, Ed. F.S. Johnson, Stanford University Press, 1965.
- Hinteregger, H. E., Representations of Solar EUV Fluxes for Aeronomical Applications, *Advances in Space Res.*, 1:39, COSPAR, 1981.
- Hoffmeister, C., Physikalische Untersuchungen an Kometen, I. Die Beziehungen des primären Schweifstrahle zum Radiusvektor, *Z. Astrophys.*, 22, 265-285, 1943.
- Holzer, T. E., and Axford, W. I., The Theory of Stellar Winds and Related Flows, *Annual Review of Astronomy and Astrophysics*, 8, pp. 31-60, 1970.
- Hudson, M. K., Lotko, W., Roth, I. and Witt, E., Solitary Waves and Double Layers on Auroral Field Lines, *J. Geophys. Res.*, 88, pp. 916-926. 1983.
- Hundhausen, A. J., Composition and Dynamics of the Solar Wind Plasma, *Rev. Geophys. Space Phys.*, 8, pp. 729, 1970.
- Hundhausen, A. J., Bame, S. J., Ashbridge, J. R., and Sydoriak, S. J., *J. Geophys. Res.*, 75, No. 25, pp. 4643-4657, 1970.
- Krankowski, D., Arnold, F. and Wieder, H., Recent Positive and Negative Ion Composition Measurements in the Lower Ionosphere by Means of Mass Spectrometers, pp 19-28 in *Magnetosphere-Ionosphere Interactions*, Ed. K. Folkestad, Universitetsforlaget, Oslo, 1971
- Krimigis, S. M., The Encounter of Voyager with Neptunes Magnetosphere, in *Magnetospheric Physics: Achievements and Prospects*, Ed. B. Hultqvist and G.-G. Fälthammar, Plenum Publ. Co. 1990.
- Lanzerotti, L., Large-Scale Organization of Solar System Plasmas, in *Magnetospheric Physics: Achievements and Prospects*, Ed. B. Hultqvist and G.-G. Fälthammar, Plenum Publ. Co. 1990.
- Lindeman, R. A., Vondrak, R. R., and Freeman, J. W., The Interaction Between an Impact Produced Gas Cloud and the Solar Wind on the Lunar Surface, *J. Geophys. Res.*, 79, pp. 2287-2296, 1974.
- Maddox, J., Down With the Big Bang, *Nature*, 340, p. 425, 1989.
- Makino, M., and Ogawa, T., An Empirical Model of Atmosperic Electrical Conductivity, *Research Letters of Atmospheric Electricity*, 4, pp. 1-4, 1984.
- Martres, M. J., and Bruzek, A., Active Regions, pp. 53-70, in *Illustrated Glossary for Solar and Solar-Terrestrial Physics*, Eds. Bruzek, A., and Durrant, C. J., D. Reidel Publ. Co., Dordrecht, Holland, 1977.

- McIlwain, C. E., Direct Measurements of Particles Producing Visible Auroras, *J. Geophys. Res.*, *65*, No. 9, pp. 2727-2747, 1960.
- McIlwain, C., Coordinates for Mapping the Distribution of Magnetically Trapped Particles, *J. Geophys. Res.*, *66*, No. 11, pp. 3681-3691, 1961.
- McIlwain, C., Magnetic Coordinates, page 45 in *Radiation Trapped in the Earth's Magnetic Field*, Ed. B. McCormac, D. Reidel Publ. Co., Dordrecht, Holland, 1966.
- Morris, M. and Rickard, L.J.: Molecular Clouds in Galaxies, *Ann. Rev. Astron. Astrophys.* *20*, pp. 517-546. 1982.
- Mozer, F. S., Cattell, C. A., Hudson, M. K., Lysak, R. L., and Torbert, R. B., Measurements and Theories of Low Altitude Auroral Particle Acceleration, *Space Sci. Rev.*, *27*, p. 155, 1980.
- Ness, N. F., Searce, C. S. and Seek, J. B., Initial Results of the IMP 1 Magnetic Field Experiment, *J. Geophys. Res.*, *69*, pp. 3531-3569, 1964.
- Ness, N. F., NASA/GSFC Report X-616-69-334, 1969.
- Ness, N. F., The Magnetic Fields of Mercury, Mars and the Moon, *Ann. Rev. Earth Planet. Sci.*, *7*, pp. 249-288, 1979.
- Ness, N. F., and Wilcox, J. M., Solar Origin of the Interplanetary Magnetic Field, *Phys. Rev. Letters*, *13*, No. 15, pp. 461-464, 1964.
- O'Brien, K., Calculated Cosmic Ray Ionization in the Lower Atmosphere, *J. Geophys. Res.* *75*, No. 22, pp. 4357-4359, 1970.
- O'Brien, K., Secular Variations in the Production of Cosmogenetic Isotopes in the Earth's Atmosphere, *J. Geophys. Res.*, *84*, p. 423, 1979.
- Pedersen, P. O., *The Propagation of Radio Waves*, Danmarks Videnskapslige Samfund 1927.
- Petelski, E. F., Fahr, H. J., Ripken, F. W., Brenning, N., and Axnäs, I., Enhanced Interaction of the Solar Wind and the Interstellar Neutral Gas by Virtue of a Critical Velocity Effect, *Astronomy and Astrophysics*, *87*, p. 20, 1980.
- Philbrick, C. R., Barnett, J., Gerndt, R., Offermann, D., Pendleton, W. Jr., Schlüter, P., Schmidlin, F., and Witt, G., Temperature Measurements During the CAMP Program, *Advances in Space Res.*, *4*, p. 152, 1984.
- Pierce, A. K. and Allen, R. G. The Solar Spectrum Between 0.3 and 10 nm, in *The Solar Output and Its Variation*, Ed. O.R. White, Colorado Associated University Press, Boulder 1977.
- Pope, J. A., A Correction to the Exospheric Electron Density Estimate Using the Nose Whistlers of March 19, 1959, *J. Geophys. Res.*, *67*, p. 412, 1962.
- Raadu, M., The Role of Electrostatic Instabilities in the Critical Ionization Velocity Mechanism, *Astrophys. Space Sci.*, *55*, p. 125, 1978.
- Sanders, D. B., Solomon, P. M. and Scoville, N. Z., Giant Molecular Clouds in the Galaxy.1. The Axisymmetric Distribution of H₂, *Astrophys. J.*, *276*, pp. 182-203, 1984.

Stone, E. C., The Voyager Mission: Encounters With Saturn, *J. Geophys. Res.*, **88**, pp. 8639-8642, 1983.

Störmer, C.: Sur un problème relatif aux mouvements des corpuscules électriques dans l'espace cosmique, *Videnskapsselskabets Skr. Mat.-nat.Kl.*, No. 14, Kristiania, Oslo, 1913

Temerin, M., Cerny, K., Lotko, W., and Mozer, F. S., Observations of Double Layers and Solitary Waves in Auroral Plasma, *Phys. Rev. Lett.*, **48**, pp. 1175-1179, 1982.

Tsyganenko, N. A., Global Quantitative Models of the Geomagnetic Field in the Cislunar Magnetosphere for Different Disturbance Levels, *Planet. Space Sci.*, **35**, pp. 1347-1358, 1987.

Tsyganenko, N.A., A Magnetospheric Magnetic Field Model With a Warped Tail Current Sheet, *Planet. Space Sci.*, **37**, pp. 5-20, 1989.

Valenzuela, A. *et al.*, The AMPTE Artificial Comet Experiment, *Nature*, **320**, pp. 700-723, 1986.

Van Allen, J. A., and Frank, L. A., Radiation Around the Earth to a Radial Distance of 107400 km, *Nature*, **183**, pp. 430-434, 1959.

Williams, D. J., The Earth's Ring Current, *Physica Scripta*, **T18**, pp. 140-151, 1987.

- Aberration 127
- Absolute magnitude 138
- Active experiments 12
- Adiabatic acceleration 96
- Adiabatic invariance 94
- Adiabatic invariant 94
- Alfvén-Mach number 128
- Alfvén velocity 10
- Alfvén waves 10
- Ambiplasma 144
- AMPTE 151
- Anomalous resistivity 82
- Antimatter 143
- Antiparticles 143
- Apparent magnitude 138
- Archimedes spiral 127
- Aarcs with ray structure 75
- Atmosphere 11
- Aurora 74
- Auroral activity 72
- Auroral oval 76
- Auroral zone 76
- Azimutal drift motion 93
- Azimutal period 94
- B-L-system 97
- Betatron acceleration 96
- Big Bang 144
- Boundary layer 65
- Bow shock 12
- Brightness 137
- Butterfly diagram 120
- Cahill discontinuity 3
- Cellular structures 11
- Centrifugal drift 91
- Chapman profile 24, 38
- Chromosphere 112
- CIV 134
- Coronagraph 113
- Collision frequency 34
- Collisionless shock wave 12
- Conductivity 49
- Conductivity tensor 29
- Convective transport 111
- Corona 113
- Coronium 113
- Cosmic radiation 12
- Cosmogony 135
- Coulomb logarithm 9
- Cowling conductivity 52
- Cosmic
- Cosmic Ray Albedo Neutron Decay 68
- CRAND 68
- Critical frequency 36
- Critical ionization velocity 134
- Cross-tail current 73
- Curtains 75
- Curvature drift 91
- Damping 57
- Debye length 8
- Debye sphere 8
- Decameter waves 104
- Dielectric tensor 55
- Differential rotation 114
- Differential spectrum 145
- Diffusion constant 132
- Diffusion equation 132
- Dip-poles 60
- Dipole field 60
- Dipole moment 85
- Drift motion 91
- Dynamic pressure 70
- Earth's radiation belts 3
- EISCAT 25
- electric double layer 82
- Electron gyro frequency 34
- Entry layer 66
- Equivalent cross section 9
- Equivalent temperature 3
- ESA 68
- ESA-Base 68
- Escape velocity 20
- ESTEC 68
- Evershed-effect 115
- Excentric dipole 60
- Exobase 20
- Exosphere 20
- Extra ordinary wave 37
- Faculae 117
- Faraday rotation 37
- Fermi acceleration 96
- Fibrilles 117
- Field line conservation 133
- Filamentary structures 11
- First adiabatic invariant 94
- Flare 118
- Flux conservation 133
- Flux invariant 95
- Forbidden regions 2
- Forbush decrease 148
- Fourth state of matter 7
- Fraunhofer lines 112
- Frozen-in magnetic field lines 10
- Galaxy 143
- Generalized Ohm's Law 49
- Geocorona 19
- Geomagnetic field 60
- Geomagnetic storm 72
- Granulation 112
- Group velocity 35
- Growth phase 73
- gyroperiod 94
- Gyro frequency 43
- Gyro center 44
- Gyro radius 9, 43
- Hall current (density) 50

- Halo 143
- Heat conduction 111
- Heliopause 12
- Heliosphere 12
- Hertzsprung–Russell diagram 139
- Heterosphere 16
- HI–regions 141
- III–regions 141
- Homogeneous arc 75
- Homosphere 16
- Hubble constant 143
- Hybrid frequency 35
- IGY 120
- Incoherent scattering 25
- Index of refraction 36
- Index of refraction 56
- Inhomogeneities 11
- Intergalactic plasma 12
- International Geophysical Year 120
- Invariant latitude 97
- Interplanetary space 12
- Interstellar plasma 12
- Ionogram 24
- Ionopause 101
- Ionosonde 24
- Ionosphere 11
- Ionospheric disturbance 32
- Ion gyro frequency 34
- Ion plasma frequency 35
- Irradiance 21
- Kennelly–Heaviside layer 24
- Kinetic temperature 14
- Koino matter 143
- L–value 97
- Longitudinal invariant 95
- Longitudinal period 93
- Longitudinal period 94
- Loss cone 96
- Lower 35
- Mach number 128
- Magnetic field lines 86
- Magnetic field vector 86
- Magnetic flux 95
- Magnetic mirrors 93
- Magnetic moment 44
- Magnetic pumping 149
- Magnetic Reynold number 132
- Magnetic storm 73
- Magnetogram 72
- Magnetohydrodynamic waves 10
- Magnetopause 61
- Magnetosheath 65
- Magnetosphere 3
- Magnetic storm 73
- Magnetospheric substorm 73
- Magnetotail 65
- Magnitude 137
- Main phase 73
- Mariner 101
- Maunder–minimum 120
- Mean free path 20
- Mesopause 20
- Mesosphere 20
- Metagalactic 143
- Milky Way 141
- Neutral sheet 65
- Neutron albedo 148
- Nose frequency 61
- Omnidirectional spectrum 146
- One–fluid model 48
- Oort’s cloud 107
- Ordinary wave 37
- Parallell conductivity 29
- parallell current 50
- Pedersen conductivity 29
- Pedersen current 50
- Pedersen profile 38
- penumbra 115
- Pfotzer maximum 67, 148
- Phase velocity 35, 57, 58
- Photographic Magnitude 137
- Photosphere 112
- Pitch angle 95
- Plage 116
- Plasma 3
- Plasmapause 63
- Plasmasheet 65
- Plasmasphere 65
- Plasma frequency 34
- Plasma mantle 66
- Plasma sheet 65
- Plasma torus 103
- Polarization 54
- Polarization field 51
- Polar plumes 65
- Prominence 117
- Proton cycle 110
- Pulsar 109
- Quasineutrality 8
- Radial diffusion 98
- Radiation belts 68
- Radiative transport 111
- Radio wave emission 104
- Rays (auroral) 75
- Recovery phase 73
- Rigidity 145
- Rosseland field 40
- Scalar potential 84
- Scale height 17
- Scintillation 37
- Second adiabatic invariant 95
- Solar activity 32
- Solar atmosphere 112

Solar constant 21
Solar cycle 14
Solar eruption 118
Solar flare 32
Solar wind 12
Sonagram 61
Sounding rocket 89
Sound speed 16
Spectral class 137
Spicules 116
Spiral galaxy 141
Spread-F 37
Stagnation point 65
Standard Atmosphere 13
Standoff distance 70
Stars 12
Stellar and interstellar plasma 12
Stellar wind 140
Stratopause 20
Stratosphere 20
Substorm 73
Sunspot 115
Sun 12
Synchrotron acceleration 149

Thermonuclear fusion 5
Thermosphere 20
Third invariant 95
Thomson scattering 25
Titius-Bode's law 134
Transport processes 19
Transverse 37
Tropopause 19
Troposphere 19
Trough 65
Turbopause 16

Umbra 115
Upper hybrid frequency 35
USSA76 13

Van Allen belts 3
Visual magnitude 137

Wave polarisation 56
whistler mode 59
Whistler waves 10
World Data Centers 72
Zeeman effect 114

- Ahay 113
 Akasofu 73
 Akasofu 79
 Albert 82
 Alfvén 2, 6, 10, 81, 82, 91, 107, 114, 124, 132, 134, 142, 143, 144, 149
 Allen 22
 Allkover 148
 Appleton 24, 53
 Aristotle 74
 Axford 70
 Axford 131
 Axnäs 136

 Balsiger 107, 108
 Barkhausen 61
 Barnett 24
 Belcher 106
 Biermann 123
 Birkeland 75
 Bolin 19
 Boström 83
 Bowyer 108
 Brenning 135, 136
 Bruzek 116, 117, 118
 Burns 144
 Burton 72, 141

 Carlqvist 120, 124
 Celsius 74
 Chiu 83
 Cloutier 103
 Craven 77, 81

 Daly 68
 Danielsson 135
 Durant 116, 117, 118

 Eather 74, 76, 78
 Evans 82

 Fahleson 135
 Fermi 149
 Folkestad 26
 Forbush 148
 Frank 67, 77, 81
 Friedman 111, 128
 Frohlich 21
 Fälthammar 2, 6, 9, 77, 82, 91, 114, 124, 125, 149

 Galeev 136
 Gibson 119
 Grieder 148

 Haerendel 150, 151, 153
 Hale 23, 114
 Hanson 28, 29, 30
 Hargreaves 39
 Hartree 53
 Heaviside 24

 Helliwell 62
 Herlofson 143
 Hinteregger 22
 Holzer 131
 Hones 80
 Hudson 83
 Hundhausen 125

 Kennelly 24
 Krankowski 137
 Krimigis 100, 104, 106

 Langmuir 7
 Lanzerotti 78, 100, 101, 102, 104, 106
 Lerner 144
 Lüst 123

 Maddox 144
 Makino 23
 Marconi 24
 McIlwain 82, 97, 98
 Mendillo 150
 Menzel 114
 Morris 141

 Ness 100, 101, 106, 125, 127
 Northrop 94

 O'Brien 148
 Ogawa 23

 Panofsky 88
 Paresce 108
 Petelski 136
 Philbrick 21
 Phillips 88
 Pierce 22

 Raadu 136
 Richard 141
 Robinson 120, 121, 143
 Ruzmaikin 142

 Sanders 141
 Schechter 143
 Seltzer 146
 Singer 67
 Soderblom 103
 Störmer 2, 67, 75
 Stewart 24
 Stone 99

 Tascione 34, 73, 119
 Temerin 83
 Tsyganenko 86

 Uberoi 101, 102, 106

 Van Allen 67
 Wallenqvist 139, 140
 Weinberg 144
 Weissman 107
 Wilcox 127
 Wyller 112

Appendix

Natural constants

Name	Notation	Value
Universal gas constant	R	$8.314 \text{ J}/(\text{mol}\cdot\text{K})$
Avogadro's number	A	$6.022 \cdot 10^{23} \text{ particles/mole}$
Boltzmann's constant	k	$1.381 \cdot 10^{-23} \text{ J/K}$
Permittivity of free space	ϵ_0	$8.854 \cdot 10^{-12} \approx \frac{10^{-9}}{36\pi} \text{ F/m}$
Electrone charge	e	$1.602 \cdot 10^{-19} \text{ C}$
Electrone mass	m_e	$9.109 \cdot 10^{-31} \text{ kg}$
Gravity constant	G	$6.673 \cdot 10^{-11} \text{ N}\cdot\text{m}^2/\text{kg}^2$
Velocity of light in vacuum	c	$2.998 \cdot 10^8 \text{ m/s}$
Permeability of free space	μ_0	$4\pi \cdot 10^{-7} \text{ H/m}$
Proton mass	m_p	$1.673 \cdot 10^{-27} \text{ kg}$

Common units

Unit	Notation	Value
Astronomic unit	AU	$1.496 \cdot 10^{11} \text{ m}$
Gamma	γ	10^{-9} T
Gauss	G	10^{-4} T
Earth's radius	R_e	$6.378 \cdot 10^6 \text{ m}$
Light-year		$9.461 \cdot 10^{15} \text{ m}$
Parsec	pc	$3.086 \cdot 10^{16} \text{ m}$

Index of notations

Symbol	Quantity	Section
a	Magnetic dipole moment	4.6
a_a	Absorption coefficient	3.5
a_j	Ionisation coefficient	3.5
a_r	Recombination coefficient	3.5
A	Avogadro's constant	2.3
b	Magnetic flux density	3.9
B	Magnetic flux density	1.3
\hat{B}	Unit vector in the direction of the magnetic field	3.9
B_p	Magnetic polar field strength	4.6
c	Velocity of light in vacuum	3.1
C_P	Specific heat at constant pressure	6.2
C_V	Specific heat at constant volume	6.2
D	Diffusion constant	6.3
D	Electric displacement	3.9
ds	Vector line element	4.6
e	Electron charge (absolute value)	1.2
e	Base of natural logarithm (2.718...)	2.3
E	Energy	9
E	Electric field	3.2
E_{eff}	Effective electric field	3.8
E_0	Energy per unit mass	6.2
E_0	Primary electric field	3.8
E_1	Secondary electric field (Polarisation field)	3.8
$E_{ }$	Parallel (to the magnetic field) electric field	3.7
E_{\perp}	Perpendicular (to the magnetic field) electric field	3.7
f	Frequency	3.1
f_{cr}	Critical frequency	3.4
f_{ge}	Electron gyro frequency	3.4
f_{gi}	Ion gyro frequency	3.4
f_{lh}	Lower hybrid frequency	3.4
f_n	Nose frequency	4.1
f_{pe}	Electron plasma frequency	3.4
f_{pi}	Ion plasma frequency	3.4
f_{uh}	Upper hybrid frequency	3.4
F	Force	3.7
\bar{F}	Average force	4.8
F_c	Centrifugal force (absolute value)	4.8
$F_{ }$	Force component along magnetic field	3.7
F_E	Electric force	4.7
F_m	Magnetic force	3.7

$F(r)$	Numerical factor	6.2
g	Acceleration of gravity	2.3
g_n^m	Multipole coefficient	4.6
G	Gravitational constant	6.2
\mathbf{h}	Magnetic field	3.9
h_n^m	Multipole coefficient	4.6
H	Scale height	2.3
H_e	Scale height of electrons	3.6
H_i	Scale height of ions	3.6
i	Imaginary unit	3.9
i_H	Hall current density	3.8
i_P	Pedersen current density	3.8
i_{tot}	Total current density	3.8
i_{\parallel}	Current density along magnetic field	3.8
\mathbf{i}	Current density vector	3.8
I	Radiation intensity	3.5
I	Light intensity	7
J	Longitudinal invariant (Second invariant)	4.8
k	Boltzmann's constant	1.2
\mathbf{k}	Wave vector	3.9
K	Numeric factor	4.4
l_c	Characteristic length	1.3
L	L-value	4.8
m	Mean molecular mass	2.2
m	Particle mass	3.7
m_e	Electron mass	3.1
m_i	Ion mass	3.4
\bar{m}_i	Local average ion mass	3.6
m_p	Proton mass	6.1
m_0	Rest mass	9
m_{\odot}	Sun's mass	6.2
M	Mach number	6.1
M	Magnitude	7
M	Molar mass	2.3
M_A	Alfvén-Mach number	6.1
n	Plasma density	1.3
n	Index of refraction	3.4
n_e	Electron density	1.3
n_i	Ion density	3.6
n_n	Neutral gas density	3.5
p	Pressure	2.3
p_d	Dynamic pressure	4.4
p_e	Electron partial pressure	3.6
p_i	Ion partial pressure	3.6
p_n	Neutral gas partial pressure	3.8
\mathbf{p}	Momentum	9
P	Rigidity	9

P_n^m	Legendre function	4.6
\mathbf{P}	Electric polarisation (Dipole moment per unit volume)	3.9
q	Electric charge	3.7
q	Ionisation rate	3.5
q_i	Ion charge	3.4
r	Geocentric distance	4.1
r	Recombination rate	3.5
r	Spherical position coordinate	4.6
\mathbf{r}	Position vector	3.9
r_e	Classical electron radius	3.1
R	Universal gas constant	2.3
R	Radius of curvature	4.8
R_e	Earth's radius	4.6
R_m	Magnetic Reynolds number	6.3
S	Cross section	6.2
S_{eq}	Equivalent collision cross section	1.3
t	Time coordinate	
T	Absolute temperature	2.2
T_a	Azimuthal period	4.8
T_e	Electron temperature	1.3
T_{ekv}	Equivalent temperature	1.2
T_g	Gyro period	4.8
T_i	Ion temperature	3.6
T_l	Longitudinal period	4.8
u	Relative velocity between coordinate systems	4.7
u	Solar wind radial velocity	6.1
\mathbf{u}	Drift velocity	3.7
\mathbf{u}_\perp	Transverse (to the magnetic field) velocity	3.7
U	Dimensionsless attenuation (URSI notation)	3.9
\mathbf{v}	Velocity	3.7
\mathbf{v}	Weighted mass velocity	3.8
v_{crit}	Critical velocity	6.4
v_E	Phase velocity of extraordinary wave mode	3.4
v_L	Phase velocity of left hand polarized wave mode	3.4
v_O	Phase velocity of ordinary wave mode	3.4
v_R	Phase velocity of right hand polarized wave mode	3.4
v_{rms}	Root mean square velocity	2.2
\bar{v}_T	Linear average velocity	2.2
v_\parallel	Velocity along magnetic field	3.7
\mathbf{v}_e	Electron velocity	3.8
\mathbf{v}_{ph}	Phase velocity	3.4
\mathbf{v}_{gr}	Group velocity	3.4
\mathbf{v}_i	Ion velocity	3.8
\mathbf{v}_n	Neutral gas velocity	3.8
\mathbf{v}_\perp	Velocity transverse to magnetic field	3.7
V_A	Alfvén velocity	1.3
V_i	Ionisation potential	6.4

V_S	Sound velocity	2.2
W	Average energy	1.2
x	Cartesian coordinate	
X	Dimensionsless frequency (URSI notation)	3.9
y	Cartesian coordinate	
\mathbf{Y}	Dimensionless vector along magnetic field (URSI notation)	3.9
z	Cartesian coordinate	
z_0	Depth of penetration	3.9
Z	Charge number	1.3
Z	Dimensionless frequency (URSI notation)	3.9
α	Angle of incidence	3.4
α	Pitch angle	4.8
α_f	Loss cone angle	4.8
β	Friction factor	3.8
β	Beta value (ratio between kinetic and magnetic pressure)	4.9
γ	Ratio between C_P and C_V	2.2
γ_n^m	Multipole coefficient	4.6
δ	Slowness parameter	4.8
ϵ_0	Permittivity of free space	1.3
θ	Spherical coordinate (polar angle)	4.6
κ_n^m	Multipole coefficient	4.6
λ	Latitude	5
λ_D	Debye length	1.3
Λ	Invariant latitude	4.8
Λ	Parameter in the Coulomb logarithm ($\ln \Lambda$)	1.3
μ	Magnetic moment (First invariant)	3.7
μ_0	Permeability of free space	1.3
ν_e	Elektron collision frequency	3.4
ν_{ei}	Effective electron-ion collision frequency	3.8
ν_{en}	Effective electron-neutral collision frequency	3.8
ν_{in}	Effective ion-neutral collision frequency	3.8
ϱ	Wave polarisation	3.9
ϱ	Gyro radius	3.7
ρ_m	Density	1.3
σ	Conductivity tensor	3.8
σ_C	Cowling conductivity	3.8
σ_e	Electron conductivity	3.8
σ_H	Hall conductivity	3.2
σ_i	Ion conductivity	3.8
σ_P	Pedersen conductivity	3.2
σ_{\parallel}	Conductivity along magnetic field	3.2
τ_e	Effective electron collision time	3.8
τ_i	Effective ion collision time	3.8
ϕ	Spherical coordinate	4.6
Φ	Magnetic flux	6.3
Φ	Scalar magnetic potential	4.6
Φ_a	Magnetic flux (Third invariant)	4.8

ξ	Dimensionless altitude coordinate	3.5
ψ	Spiral angle	6.1
Ψ	Scalar magnetic potential	4.6
ω	Angular velocity vector	4.7
ω	Angular frequency	3.4
ω_g	Gyro angular frequency	3.7
ω_{ge}	Electron gyro angular frequency	3.4
ω_{gi}	Ion gyro angular frequency	3.4
ω_{lh}	Lower hybrid angular frequency	3.4
ω_{pe}	Electron plasma angular frequency	3.4
ω_{pi}	Ion plasma angular frequency	3.4
ω_{\odot}	Sun's angular rotation velocity	6.1
ω_{uh}	Upper hybrid angular frequency	3.4
Ω_0	Phase rotation per unit path length	3.4