COSMOGONY AS AN EXTRAPOLATION OF MAGNETOSPHERIC RESEARCH

HANNES ALFVÉN

Royal Institute of Technology, Department of Plasma Physics, Stockholm, Sweden

and

University of California, San Diego, La Jolla, CA 92093, U.S.A.

Abstract. A theory of the origin and evolution of the Solar System (Alfvén and Arrhenius, 1975, 1976) which considered electromagnetic forces and plasma effects is revised in the light of new information supplied by space research. In situ measurements in the magnetospheres and solar wind have changed our views of basic properties of cosmic plasmas. These results can be extrapolated both *outwards in space*, to interstellar clouds, and *backwards in time*, to the formation of the solar system. The first extrapolation leads to a revision of some cloud properties which are essential for the early phases in the formation of stars and solar nebulae. The latter extrapolation makes possible to approach the cosmogonic processes by extrapolation of (rather) well-known magnetospheric phenomena.

Pioneer-Voyager observations of the Saturnian rings indicate that essential parts of their structure are 'fossils' from cosmogonic times. By using detailed information from these space missions, it seems possible to reconstruct certain events 4–5 billion years ago with an accuracy of a few percent. This will cause a change in our views of the evolution of the solar system.

Table of Contents

1. Introduction

This paper is a brief review of work on cosmogony (evolutionary history of the solar system) which began in 1942. The new idea was that electromagnetic (or hydromagnetic) effects were of decisive importance for understanding how the solar system got to its

present state. Because previous cosmogonies since Laplace considered mechanical forces alone, this was not reconcilable with the generally accepted types of cosmogonies. Certainly, these have changed drastically during the ages, but almost all of them neglected hydromagnetic and plasma effects. Few cosmogonists had more than a superficial knowledge of hydromagnetics and plasma physics, with the result that the decisive importance of the 2:3 contraction and the band structure have not been appreciated.

Space research has now changed the situation by giving us new information about electromagnetic and plasma effects in space. From *in situ* measurements in the magnetospheres we know the properties of plasmas over five or ten orders of magnitude in density, in magnetization, in temperature, etc. and we also begin to understand what processes are possible and which are not. This has introduced or is introducing a new climate in cosmical physics which may be more favorable for a serious discussion about the evolutionary history of the solar system.

1.1. New paradigm

The foundations of a space age cosmic plasma physics, which now must be slowly built up, are essentially the following:

(a) The electromagnetic spectrum can now be observed outside the atmosphere, which means that the number of octaves available has increased by more than a factor of two.

In the field of *plasma physics* we have especially to note that large parts of the newly discovered astrophysical phenomena – for example, in X-ray and gamma-ray astronomy – are obviously due to plasma phenomena. Still more important are

(b) *In situ* measurements in the magnetospheres (including the heliosphere).

(c) Laboratory studies of phenomena of interest in cosmic plasma physics.

(d) Increased understanding of how to extrapolate results obtained in one field to other fields of plasma physics.

A survey of some of the 'paradigm transitions' which this has caused or is causing has been published in a monograph (Alfvén, 1981a). Summaries of this have been presented in Alfvén (1982, 1983a c).

The following table (Table I, essentially the same as published in a *Geophysics Research* letter (Alfvén, 1983a)) is a catalogue of the fields which are up for revision.

1.2. Consequences for cosmogony

For cosmogony this has led – or is inevitably leading – to a new approach. We do not need to base cosmogonic theory on more or less reasonable *assumptions* about conditions at the time when the solar system was formed (probably 4–5 G years ago), or on uncertain interpretations of distant, marginally observable phenomena. We can instead treat cosmogony as an *extrapolation* of reasonably well-established processes from *space research*, often derived from *in situ* measurements (see Figure 1). The result is an approach in which the evolutionary history is decided by a *combination of mechanical effects and electromagnetic (plasma) effects.*
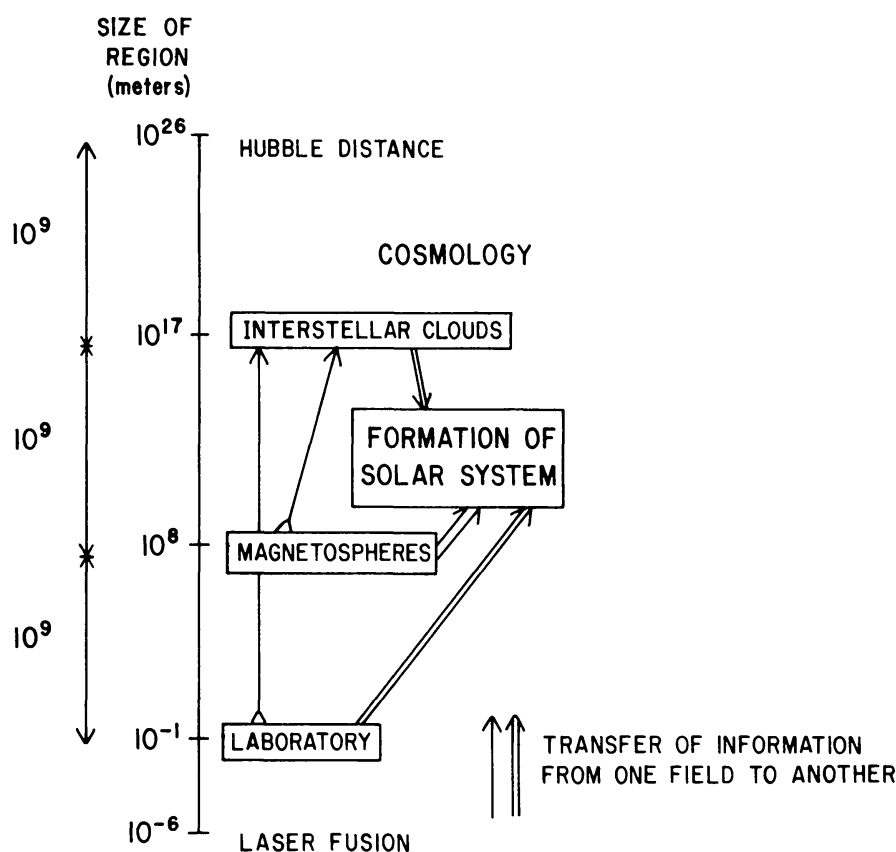
IMPORTANT FIELDS OF PLASMA PHYSICS

Fig. 1. Magnetospheric research has matured to such an extent that it is possible to treat essential parts of the evolutionary history of the solar system as an *extrapolation of magnetospheric research*. Laboratory experiments also form an important basis for this. Further, extrapolation from both magnetospheric and laboratory results contribute to a revision of our view of interstellar clouds, and hence influence also the way in which we approach cosmogony. The transfer of information from one field to another is shown in the figure.

An attempt to introduce electromagnetic effects in cosmogony was actually started long before the space age. Some of the results of investigations of this kind seem to be in agreement with the new results. For this reason a brief summary of the historical background of this approach is given in Table II, which could also be considered as a very condensed abstract of this paper. The meaning of the terms will be explained in later sections of this paper.

2. State of Cosmogony in Light of Space Research

What has been summarized in Tables I and II leads to the evolutionary history of the solar system, which in some essential respects is similar to what is described in the Alfvén–Arrhenius monographs. This is not unexpected, because, as seen in Table II, the evolution of these theories has from the beginning been coordinated with the develop-

TABLE I

Magnetospheric research

is causing a paradigm transition in geophysics and astrophysics for the following reasons:

No. 1. *Electric double layers* are realized to be very important.

No. 2. The often misleading 'magnetic merging' theories of energy transfer should be replaced by an electric current description, including the circuits in which the currents flow.

No. 3. Homogeneous models often are found to be misleading and should be extensively replaced by inhomogeneous models.

No. 4. It is realized that inhomogeneities are produced by *filamentary currents*

No. 5. and by *surface currents*, dividing space into *cells*.

No. 6. It is concluded that space in general has a *cellular structure*.

No. 7. The introduction of the current-circuit description makes it impossible to neglect the pinch effect term in the pressure equation $\nabla(p + B^2/2\mu_0) - (B\nabla)B/\mu_0 = 0$.

No. 8. It is doubtful whether large-scale *turbulence* is of importance in diffuse media.

No. 9. In a space plasma, electric currents may produce chemical separation.

No. 10. In dusty plasma, gravito-electromagnetic effects are often important.

No. 11. The 'critical velocity' is often decisive to the interaction of neutral gas and magnetized plasma. Cosmological consequences will not be discussed here.

TABLE II

History of the solar system with electromagnetic effects introduced

1942	It was shown that a cosmogonic model of this type required that *three mechanisms* be postulated: a. Electromagnetic transfer of angular momentum; b. The existence of a phenomenon which later was called 'critical velocity'; c. A plasma-planetesimal transition associated with a 2 : 3 contraction. This produced 'cosmogonic shadows'.
1954	A survey of the theory was published as a monograph *On the Origin of the Solar System* (Alfvén, 1954). It included a development of the theory of Saturnian rings leading to a correction in the contraction factor by a few percent.
1960	Laboratory confirmation of *critical velocity* (Fahleson, 1961), essentially based on a technique developed by Bratenahl (Anderson *et al.*, 1959).
1974	Zmuda and Armstrong (1974) (and Iijima and Potemra, 1978; for a survey, see Potemra, 1979) map the magnetospheric current system which gives the needed *transfer of angular velocity*.
1975, 1976	Cosmogonic theory systematically developed in two monographs by Alfvén and Arrhenius (1975, 1976).
1980	Space research calls for a 'paradigm transition'. A brief review is given in Alfvén (1981), and further developed in (1982, 1983) (summarized in Table I).
1982	*Critical velocity* effect in space demonstrated by space experiment by Haerendel (1982).
1983	Holberg's treatments of Voyager results (Holberg *et al.* 1982) make possible a further *confirmation of* 2 : 3 *contraction and cosmogonic shadow effect*. This is also supported by earlier investigations of the asteroidal belt. All together the 2 : 3 fall-down ratio is found in *seven* cases, four in the Saturnian ring an three in the asteroidal belt. This is an encouragement to the further pursuit of this cosmogonic approach.

ment in cosmic plasma physics (Alfvén, 1981a). A brief survey will be given here (compare Table III).

2.1. ELECTRIC CURRENTS IN INTERSTELLAR CLOUDS

There are good reasons for the general view that stars and solar systems are born out of an interstellar *cloud of dusty plasma*. However, the theory of the origin and evolution of such clouds and the formation of stars and solar nebula is a field which must now be revised for the following reasons:

In situ measurements in magnetospheric plasmas (including the solar wind) have caused drastic changes in our views of the properties of cosmic plasmas. What was considered sacrosanct ten or even five years ago is now hopelessly obsolete. This theoretical *paradigm transition*, which is summarized in Table I, has penetrated *as far out as in situ measurements are made*; i.e., as far as spacecraft have travelled. Outside this limit the paradigm transition has not yet taken place. Plasmas in interstellar space are still being treated according to the old paradigm. This means in reality that the present theories of interstellar clouds and of the formation of stars and solar nebulae are based on the *tacit assumption that the basic properties of cosmic plasmas change at the outer reach of spacecraft.*

It is obvious that astrophysics cannot remain in this unstable state (indeed, a 'universal instability' in plasma physics!). The new paradigm will sooner or later be extended to interstellar space. It will cause a revolutionary change in our view of the evolution of interstellar clouds, in the following respects.

(a) According to No. 2 in Table I, cosmic plasmas cannot be described by the magnetic field picture alone. This must be supplemented by an electric current description. Astrophysicists are often reluctant to accept the existence and importance of electric currents in interstellar space, but none of them claims that the magnetic fields are curl-free. As a non-curlfree magnetic field means electric currents, they *implicitly* accept that interstellar space is penetrated by electric currents. However, there is an immense difference between an implicit acceptance and an *explicit* description of the phenomena in terms of electric currents. The latter description calls immediately for *models of the circuits* in which the currents flow, and models of the *dynamos which produce the currents*. Such currents may transfer energy from one region to another, sometimes over distances comparable to the size of the whole galaxy. (With regard to the circuit description, it has been objected that 'there are no wires in space'. But 'circuits' do not necessarily mean an aggregate of simple linear elements. Especially in the 'computer age', circuits often contain non-linear distributed elements, e.g., as given in Boström's (1974) circuit of a magnetic substorm.)

(b) As soon as electric currents are introduced explicitly, attention is focussed on the *pinch effect*. In the pressure equation

$$\nabla(p + B^2/2\mu_0) - (\mathbf{B}\nabla)\mathbf{B}/\mu_0 = 0$$

the second term represents the pinch effect. If this is neglected, the sum Σp of gas pressure p and the magnetostatic pressure $B^2/2\mu_0$ should be constant. In astrophysics

there seems to be a general belief that this is usually the case. As soon as we accept that there are currents in space, this is not valid. In a typical Bennett pinch *both the pressure and the magnetic field are large inside the pinch but zero outside*. A typical simple Bennett pinch is produced when

$$\frac{\mu_0}{4\pi} I_z^2 > 2Nk(T_e + T_i)$$

(I_z = current, T_e, T_i are electron and ion temperatures, N = number of particles per unit length.) Figure 2 illustrates three typical cases of stationary and cylindrically symmetric currents (i) and magnetic field line (B) configurations. In most treatments of the evolution of an interstellar gas cloud, it is assumed that electromagnetic forces oppose the contraction, as in (a), whereas they just as well may assist or cause the contraction, as in (c). The intermediate case (b) may be a first approximation of a model of filamentary currents (e.g., see Alfvén, 1981a, p. 95).
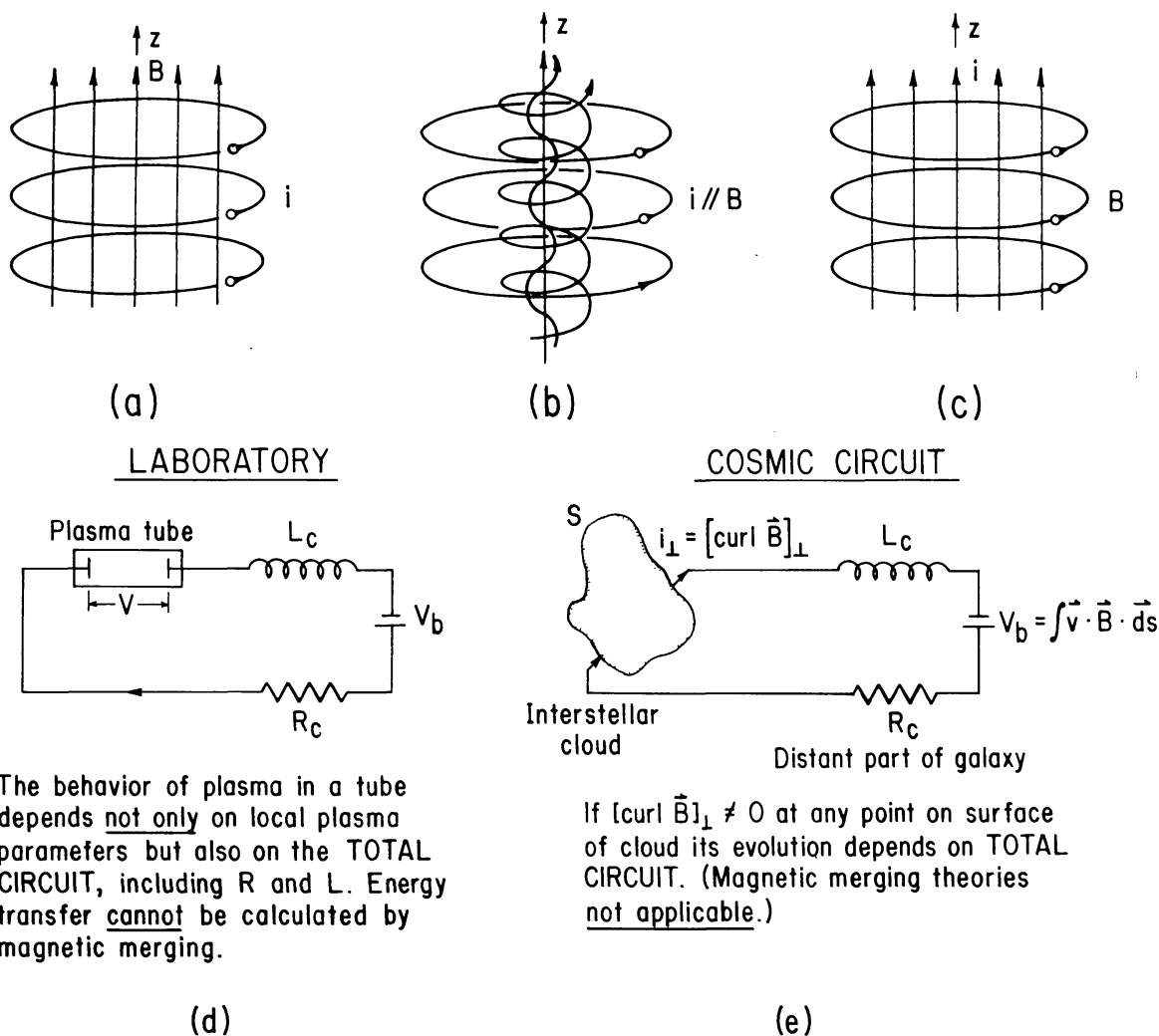

(a)   (b)   (c)

LABORATORY   COSMIC CIRCUIT

The behavior of plasma in a tube depends <u>not only</u> on local plasma parameters but also on the TOTAL CIRCUIT, including R and L. Energy transfer <u>cannot</u> be calculated by magnetic merging.

If [curl $\vec{B}$]$_\perp \neq 0$ at any point on surface of cloud its evolution depends on TOTAL CIRCUIT. (Magnetic merging theories <u>not applicable</u>.)

(d)   (e)

Fig. 2.   Three special cases of stationary and cylindrically symmetric current (i) and magnetic field (B) configurations. (a) A toroidal current and an axial magnetic field leading to a force opposing contraction. (b) A force-free configuration with i and B parallel. (c) The Bennett pinch with an axial current and a toroidal magnetic field. Electromagnetic effects aid and even start contraction.

(c) According to No. 3, homogeneous models of plasmas are now increasingly replaced by inhomogeneous models. When a new field is opened, it is natural to approach it by making homogeneous models, in the belief that these will in any case be a reasonable first-order approximation to a final theory. In plasma physics we have the sad experience that this is very often not true. When a field has matured to such an extent that it is obvious that homogeneous models are no longer sufficient, it is often evident that inhomogeneous models give a drastically different description of the phenomena. The homogeneous model was of no use. Instead, it led the modeling into a dead-end from which it often is very difficult to turn back because a powerful establishment committed to the homogeneous model has already been formed. Dessler (1984) has drawn attention to one of many cases when such an establishment has delayed progress by decades.

(d) According to No. 4, there is often an association between electric currents and observed filaments. Examples of this in our close vicinity are auroral rays (probably) associated with filamentary currents, the filamentary structure of the solar corona, and the filamentary currents in the ionosphere of Venus (Elphic and Russell, 1983). In interstellar clouds, there are often observed filamentary structures (especially in contrast-enhanced photographs), Such observations support our conclusion that interstellar space, and not the least interstellar clouds, are penetrated by a network of electric currents (cf. Alfvén, 1981a, Chapter II).

Concerning clouds in which no filamentary structure is observed, it is an open question whether this depends on an absence of them or the inadequacy of observational methods to detect them. From the general picture of the new paradigm the latter interpretation seems to be preferable.

(e) No. 5 and 6: It is not obvious that these are decisive for the present development of cosmogony.

(f) No. 7: This has already been discussed in the beginning of this section.

(g) No. 8: Turbulence is generally believed to be decisive for the evolution of interstellar clouds and the formation of the solar system. There seems to be no convincing observational evidence for this (see Alfvén, 1981a, p. 84).

2.2. New approach to the evolution of interstellar clouds

As stated above, sooner or later the new paradigm will penetrate also the field of the evolution of interstellar clouds. The theory of interstellar clouds should be treated as an *extrapolation of magnetospheric research* (Alfvén, 1981a) (see Figure 1).

Very much work will be required for this transaction, and it is difficult to predict in detail what the result will be. As a reasonable guess as to what a future model of the formation and evolution of interstellar clouds should be, we may suggest the following:

(a) Electric currents in 'void' interstellar space assist gravitation in collecting matter by the pinch effect, so that interstellar clouds are formed.

(b) These develop under the combined action of mechanical and electromagnetic forces. The volume occupied by currents may constitute a very small fraction of the total volume, so that the plasma regions are not evident in the averages of measurements with

insufficient resolution. Still, a network of filamentary currents may be decisive to the evolution of the clouds. It is correct to treat the evolution of an interstellar cloud independent of its surroundings *only* if there is no current connecting it with the surroundings (cf. Figure 2e).

(c) As stated above (compare Figure 2), the general belief that electromagnetic forces oppose the contraction of a cloud is not necessarily correct. Pinch effects may contribute to the contraction and, indeed, cause a collapse of clouds with a mass that is orders of magnitude smaller than the Jeans mass.

(d) A 'stellesimal' star formation out of a dusty cloud seems possible (compare Alfvén, 1981a, Chapter V).

2.3. PROPERTIES OF THE SOLAR NEBULA

When the Sun is formed it will be surrounded by a dusty plasma penetrated by a network of currents which partly support it (Figure 3). This 'solar nebula' is drastically different from the Laplacean nebula. It is possible that Oort's cometary cloud is a relic of this. The cloud is strongly inhomogeneous and contains regions of different chemical composition (compre No. 9). From this primeval cloud there rains cloudlets of different
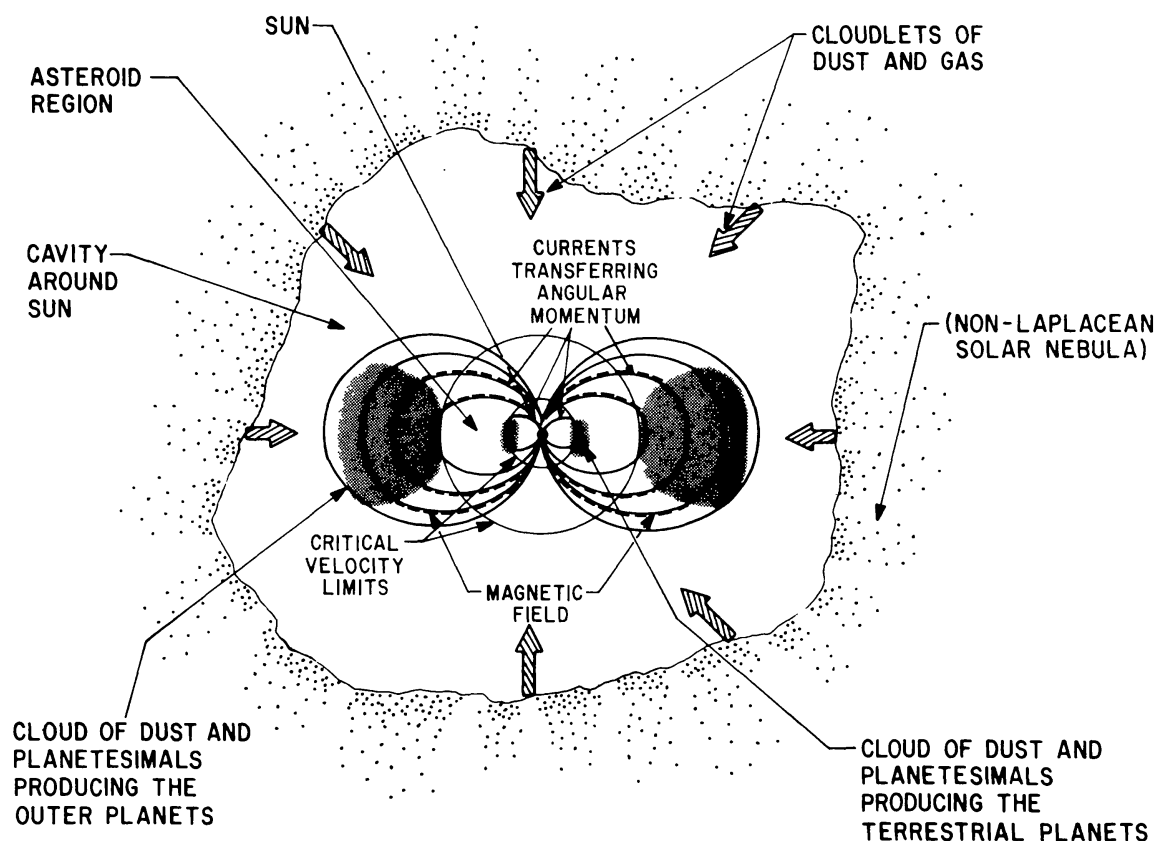

Fig. 3.   After the formation of the Sun and its magnetic field, remnants of the (non-Laplacean) solar nebula (formed according to Sections 2.2 and 2.3) fall towards the Sun in cloudlets of dust and gas. When a falling cloudlet reaches its critical velocity, it is stopped. Currents from the spinning Sun support the cloudlet for a short period of time and transfer angular momentum to it. At the plasma-planetesimal transition, grains in Kepler orbit are produced and plasma processes are then no longer important.

composition down towards the Sun. Moreover, there is a rain of cosmic dust, perhaps similar to the present rain of meteoroids from the cometary cloud. For possible models, see Alfvén and Arrhenius (1975, p. 55; 1976, Chapter 19).

From now on our attempt to reconstruct the evolutionary history of the solar system enters a new phase in two respects:

(a) Up to this point we have discussed the evolution of an interstellar cloud as a unity, even if the cloud is very inhomogeneous. From now on we have to discuss the *evolution of individual cloudlets* as distinct from the *evolution of the whole solar nebula*. The latter process is an integration of the processing of cloudlets, and consists essentially of a slow transformation of the solar nebula into planetesimals (and later planets) during a long period of time.

(b) The second reason why we are entering another phase is analytical. The process of planet formation around the Sun is similar to the formation of satellites around some planets, especially Jupiter, Saturn, and Uranus, which have well-developed satellite systems. Hence we should aim at a general theory of formation of secondary bodies around a central gravitating, rotating and magnetized body. This requirement has been referred to as the 'hetegonic principle' (see Alfvén and Arrhenius, 1975, 1976. The arguments for this view are discussed in some detail there.). The principle is related to what now is usually called 'comparative planetology', but should include also the formation of bodies around the Sun. Galileo, when discovering the Jovian satellites, already called this system a 'solar system in miniature'.

As we have four well-developed systems to base our conclusions upon, we can speak with more confidence about the processes than for the earlier states of development. Hence, our model for all four systems should start from the assumption that a magnetized central body was formed already and surrounded by a dusty plasma from which cloudlets of different chemical composition, together with dust grains, fall in towards it.

2.4. Basic processes in evolution of solar nebula

In this state there are three processes which were decisive for the present structure of the solar system: (cf. Table II).

(a) *The transfer of angular momentum* from the central body to the surrounding plasma. The transferred angular momentum is now found in the orbital moment of the secondary bodies.

There is a rather obvious candidate for this process, viz., the auroral current system, which is known to transfer angular momentum between a rotating central body and a surrounding plasma (Figure 4). Essentially the same process is well known from the Jupiter-Io system (Hill *et al.*, 1983). One of the processes which is claimed to account for the loss of momentum from the Sun is turbulence, but with reference to No. 8, and especially what will be demonstrated later, this is not an acceptable cosmogonic process.

Another process to account for the loss of solar momentum is the solar wind. This is an interesting and perhaps partially correct suggestion. However, it is not clear how it can incorporate the band structure and the cosmogonic shadow effects.
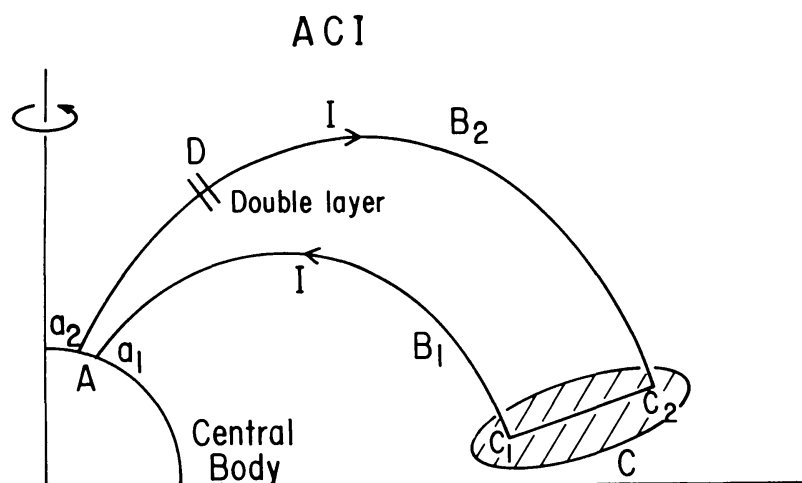
ACI

Fig. 4.   Simplified picture of the auroral circuit transferring angular omentum between a magnetized rotating body A and a surrounding plasma cloud C. Current I may produce double layers D.

With our present knowledge there seems to be no serious objection to accepting this electromagnetic transfer *qualitatively* as the basic process. When we come to a *quantitative* evaluation of the process there seems to be no serious objection to the view that the satellite systems were formed by such a transfer from the mother planet. However, for the planetary system the situation looks more difficult.

If we assumed that the formation of the solar system was a very rapid chaotic process with a time constant of less than a million years, we would run into difficulties. The magnetic field must first support all the matter and then transfer angular momentum to it. This would require an enormously strong solar magnetic field. We can avoid such a difficulty by assuming that *the planetary system was formed by cloudlets* going through the momentum transfer process with *a short time constant* T_c while this processing led to a *slow buildup* of planetesimals and planet with a time constant T_i which may be many orders of magnetude larger than T_c. The result is that the density ρ of the plasma supported at a certain moment by the magnetic field need only be a small fraction of the total smeared-out density ρ_s of the produced planetesimals: $\rho = \rho_s T_c / T_i$.

The distinction between the rapid processing of cloudlets and the slow integrated buildup means that ρ was so small that the momentum transfer processes could take place in a low-density plasma (collisionless plasma). This is fortunate because much of the study of magnetospheric plasmas has been and is concentrated on such plasmas. Hence we should be able to treat *the transition from plasma to planetesimals as an extrapolation of present-day magnetospheric results* (for details, see Alfvén and Arrhenius, 1976).

(b) *Band Structure of the Solar System.* The second basic plasma process is *the critical velocity*.

When developing a tentative early theory of the possible importance of electromagnetic processes in solar system evolution, it was necessary to postulate the existence of 'the critical velocity' in order to explain the band structure of the solar system (see

Table II). Such a process was unknown at that time, but the cosmogonic evidence for its existence was considered so compelling that laboratory experiments to demonstrate it were started as soon as possible. These were successful, and there exists now a literature of some hundred papers regarding this phenomenon (Axnäs *et al.*, 1982).

The existence of such a process in space is now confirmed by space experiments (Haerendel, 1982). As there are very few phenomena which have been predicted from a cosmogonic theory and later discovered this gives some confidence that the critical velocity is the process which is responsible for the band structure.

However, in spite of all this, the problem of planetary formation remains difficult. The critical velocity is a phenomenon in pure gases, but how a dusty plasma behaves is not clear. The most serious problem is to understand how the planets and satellites have acquired their present chemical compostion. Possible chains of processes have been discussed but a convincing solution has not yet been found.

(c) Concerning the 2:3 contraction see 3.3.

2.5. SCENARIO OF THE COSMOGONIC PROCESS

The general scenario of the cosmogonic processes is shown in Table III. Plasma effects were of considerable importance for the evolutionary history of the solar system from the formation and evolution of cosmic clouds to the formation of the Sun and a surrounding solar nebula.

(a) In the *solar nebula* the plasma effects were of decisive importance in two respects: (1) They *transferred angular momentum* from the Sun to the plasma out of which later the planets (and asteroids) were formed, and (2) The *critical velocity* produced the *band structure* of the solar system. The basic plasma processes which cause the critical velocity are still not very well clarified theoretically, but the phenomenon is extensively studied in the laboratory, and space experiments have demonstrated its importance for cosmic plasma physics.

(b) After the plasma phase of the solar nebula came the *plasma planetesimal transition* (PPT). This was not a sudden and violent turbulent transition, but a slow, continuous process working for a very long time, perhaps 10–100 million years, which continuously transformed in-falling matter in a dusty plama state into a planetesimal state. (However, the processing of individual cloudlets was a rapid process; see Section 2.4(a). This process worked not only in the planetary system, but at a later period (when the planets were formed) it worked around the planets, producing 'satellitesimals'.

(c) The mass of matter in the planetesimal state increased slowly, until the planetesimals began to *aggregate to planets*. Later, similar processes led the 'satellitesimals' to aggregate to satellites. Plasma processes are of negligible importance for these processes.

In one region in the planetary system, viz., the asteriodal region, and in one region in the satellite systems, viz., the Saturnian ring, the aggregation to planets or satellites has not taken place. The reasons for this are the low density in the asteriod region and the location of the Saturnian ring inside the Roche limit (Alfvén and Arrhenius, 1975, Chapter 13; 1976, Chapter 18). Hence, at least in certain respects, the state in these

TABLE III

Formation of planets/satellites from interstellar clouds

State of matter which is located at present in planets/satellites	Evolutionary process	Main evolutionary mechanism	References
Dusty plasma	Evolution of interstellar cloud formation of Sun and solar nebula	Gravitation Pinch effect	2.1, 2.2, 2.3
	Evolution of solar nebula	Electro-magnetic transfer of angular momentum Critical velocity	2.4 2.5
Planetesimals	Plasma-planetesimal transition	2/3 contraction cosmogonic shadow effect Rosseland field	3.3
	Accretion of planetesimals to planets	Mechanical effects plasma processes not important	
Planets Satellites	Formation of satellites around planets occurs by a repetition in miniature of these processes (starting with formation of nebula around planet).		

TABLE IV

Cosmogonic shadows

Saturnian ring from Holberg's data

Mimas	3.075	
Co-orbitals	2.510	0.646
Shepherds	2.349	0.63
	2.310	
Cassini Center	1.984	0.655 (0.650–0.660)
Outer B	1.945	
Holberg min	1.58	
Inner B	1.525	0.635
Inner C	1.235	

Average 0.642 ± 2%

Asteroidal region

Jupiter	5.18	0.676
Main belt outer limit	3.50	
High density outer limit	3.22	0.674
High density inner limit	2.36	0.683
Main belt inner limit	2.20	
Theoretical value		
		0.667

(Alfvén, 1981b).

regions represents the planetesimal state. This makes them of decisive importance for our attempts to reconstruct the plasma-planetesimal transition. Because much – or rather most – of the cosmogonic information initially stored in the planetesimal state is obliterated in the planetesimal-planet transition, they are of unique value for clarifying the evolutionary history of the solar system.

In the following, our approach to cosmogony is to a large extent based on a study of the plasma-planetesimal transition. It turns out that the Saturnian ring and the asteroidal belt contain information which is decisive for our reconstruction of the history of the solar system. In this paper we concentrate our attention on the Saturnian ring.

2.6. STRUCTURE OF THE SATURNIAN RING ACCORDING TO THE PIONEER-VOYAGER OBSERVATIONS

The Pioneer and Voyager exploration of the Saturnian rings has given us most valuable material about its structure. For reasons given in Alfvén (1983c), we concentrate our attention here on the bulk structure, which seems to give information of decisive value for clarifying the origin of the rings. That paper relied on Holberg's curves (Holberg *et al.*, 1982; Holberg, 1983). Now similar curves by Esposito *et al.* (1983) have also been published. In the points that are of interest to us the newly published curves agree very well with those we have used.

3. Scenario of the Saturnian Ring Formation

What has been said in Section 2.5 leads to the following *scenario for the formation of the ring system* (see Alfvén and Arrhenius, 1975, 1976; Alfvén, 1983b, c).

(1) Saturn was already formed, with approximately its present mass and present spin. Its magnetic field may also have had the present shape (close to a dipole field) but we do not know its strength.

(2) Cloudlets of gas and dust from interplanetary space fell in towards Saturn. They became ionized, which led to currents of the same type as mapped in the planetary magnetosphere by Zmuda and Armstrong (1974) and later by Iijima and Potemra (1978). (The Jupiter-Io circuit is similar (see Section 2.4(a)). This current system transferred angular momentum from the planet to the plasma (Alfvén and Arrhenius, 1976, Chapter 16; Alfvén, 1981a; pp. 52, 120).

(3) This brought the plasma into a state of partial corotation, so that Saturn's gravitation was compensated to *two-thirds* by the centrifugal force and to *one-third* by electromagnetic forces (from the Saturnian magnetic field; see Figure 5).

(4) At the transition *from the plasma to the planetesimal phase* the electromagnetic forces vanished, which caused a *contraction by a factor* $\Gamma = \frac{2}{3}$. (This factor is given by the geometry of Saturn's magnetic dipole field; see Figure 5.) Early Voyager results have already demonstrated that there is strong evidence for this process in the present structure of the Saturnian ring (Alfvén, 1983b).

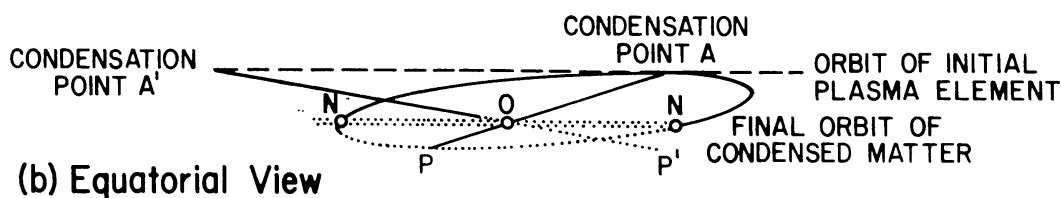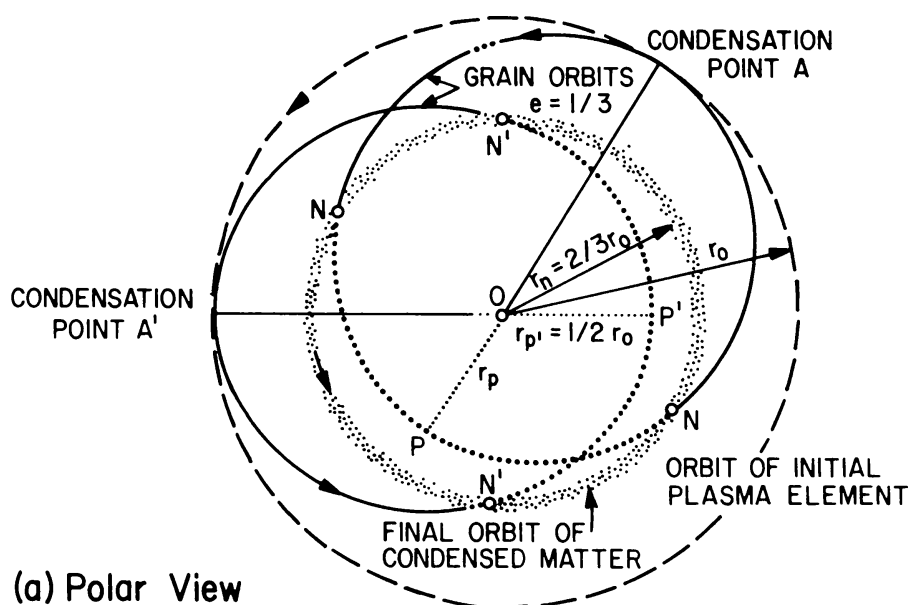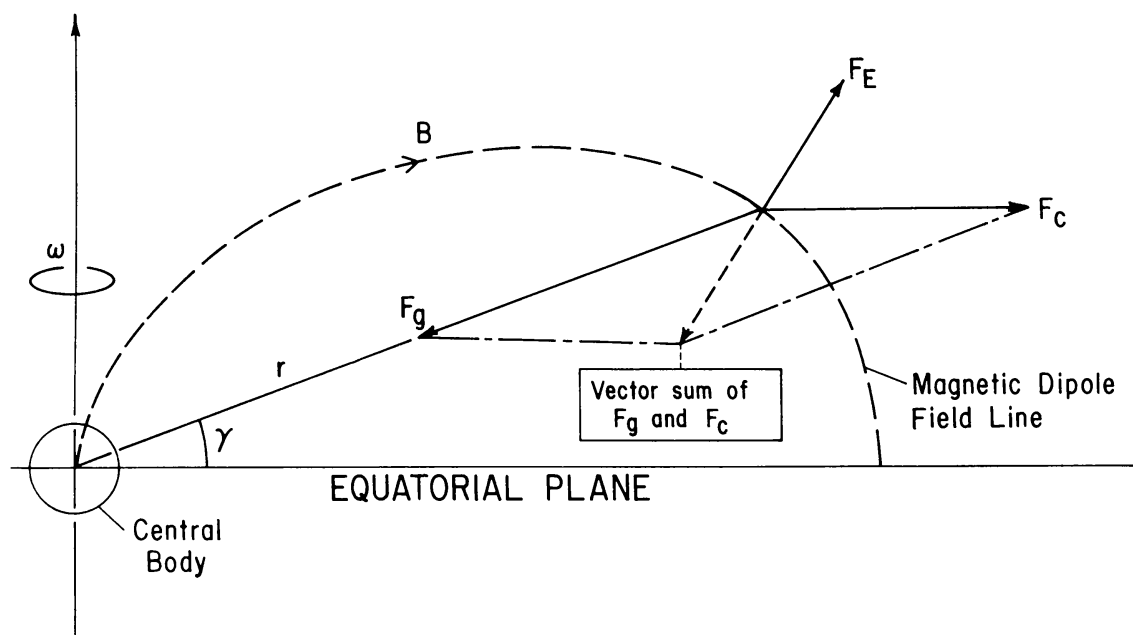(a) Polar View

(b) Equatorial View

Fig. 5. (a) Charged particles (plasma, charged dust, or cloudlets) in an axisymmetric magnetic dipole field around a gravitating rotating body. If their motion is magnetic-field dominated, a quasi-stationary motion requires that the vector sum of gravitation and centrifugal force be perpendicular to the magnetic field line. As shown by Alfvén and Arrhenius (1975, 1976), this means $v_\theta = (\frac{2}{3})^{1/2} v_K$, where v_θ is the rotational velocity and v_K is the Kepler velocity. (b) Vanishing magnetic forces give a transfer into elliptic orbits. If the magnetic field or the particle charge suddenly disappears, the particles at the central distance a_0 will orbit in ellipses with semi-major axis $a = \frac{3}{4}a_0$, and eccentricity $e = \frac{1}{3}$. They will collide mutually when they reach the nodes in the equatorial plane with $a = \frac{2}{3}a_0$, and after collisions move in a circle at $a = \Gamma a_0$ with $\Gamma = 2:3$.

3.1. Negative diffusion and stability

The first question we have to answer must be: Is it reasonable that essential parts of the present ring structure are a *'fossil' from cosmogonic times* (4–5 billion years ago)?

Baxter and Thompson (1971, 1973) have demonstrated that under certain conditions the *diffusion* in a population of grains in Kepler orbits is *'negative'* (see Figure 6). This

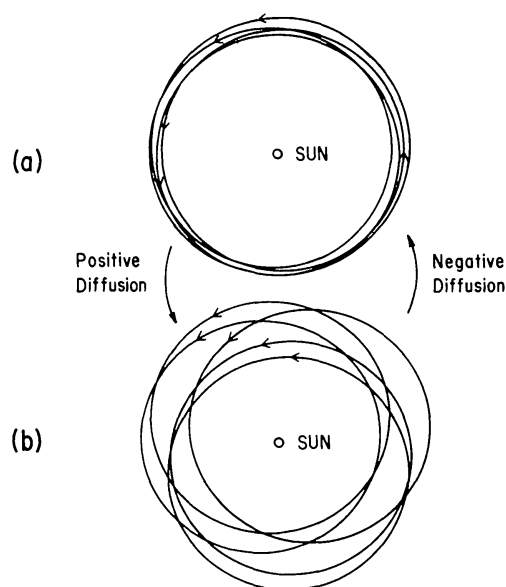EVOLUTION OF BODIES MOVING IN KEPLER
ORBITS AND INTERACTING

Fig. 6. Interaction of a large number of particles in Kepler orbits. In the discussion of collisions between particles in interplanetary space (e.g., evolution of the asteroidal belt or meteor streams) it is usually taken for granted that state *a* will evolve into state *b* (positive diffusion). This is usually *not* correct. Collisions between the particles will *not* spread the orbits since the diffusion coefficient is negative (Baxter and Tompson, 1971, 1973). Instead, collisions will lead to equalization of the orbital elements, leading from state *b* to state *c* so that a jet stream is formed.

result is confirmed by Lin and Bodenheimer (1981). The present ring system consists of 1000, if not 10 000, ringlets (cf. Figure 7). This indicates that the negative diffusion mechanism is active today. There seems to be no obvious objection to the assumption that the same mechanism was active in the past. This means that the present structure may derive from cosmogonic times (Alfvén, 1983c). Hence, it is meaningful to try to reconstruct essential events in the evolutionary history from 'fossils' stored in the Saturnian ring. There are reasons to believe that this holds also for the asteroidal belt.

3.2. Non-catastrophic formation

The general scenario must be, consequently, that the Saturnian satellites and ring system were formed *not* as a result of a sudden event, but by a slow injection of diffuse matter during a period of millions of years. Very early during this period a ring-satellite system was already formed which was qualitatively similar to the present one, but with only a
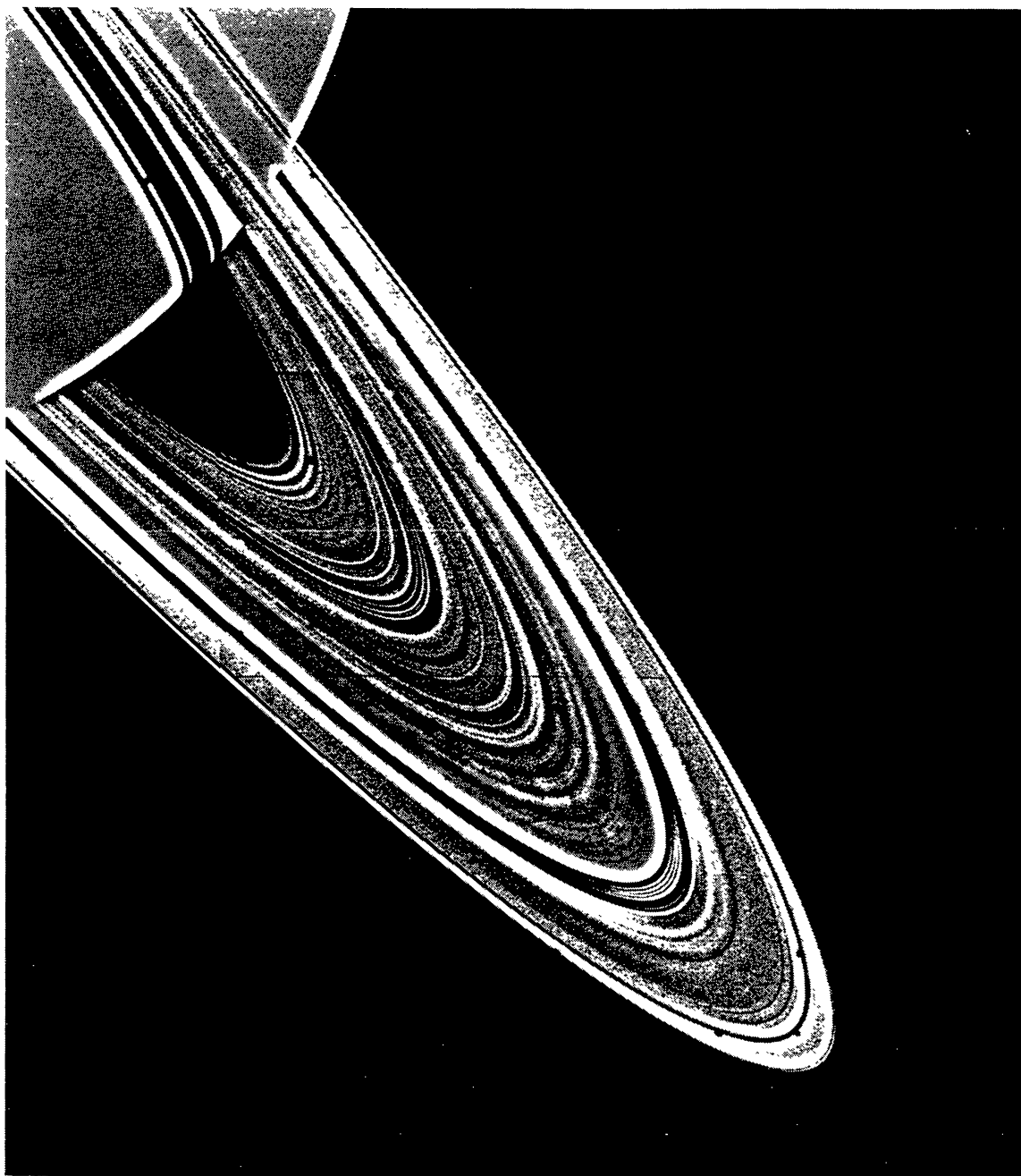
Fig. 7.   This photo of the Saturnian Ring shows some of its fine structure. In order to maintain a large number of ringlets, a negative diffusion process is indicated.

small fraction of its present mass. When more mass accumulated the bodies became more massive, but the same structure was retained. Hence, during most of the time of accretion shadow-producing bodies were located at the same places as today.

3.3. COSMOGONIC SHADOW

In the cosmogonic model we discuss it was assumed that in the Saturnian magnetosphere there was a dusty plasma which, to some extent, was concentrated at the equatorial

plane. At the Saturnian distance of Mimas, this satellite (or the jet stream out of which it was formed) swept the plasma, so that a 'hole' – in reality an empty ring – was produced.

During the Pioneer mission, Fillius and McIlwain (1980) actually observed this kind of phenomenon. In fact, at the distance of Mimas, the observed plasma density went down by orders of magnitude (see Figure 8). Janus produced a similar although more narrow and shallower 'hole'. The shepherd satellites and the A ring gave an almost complete cut-off in the plasma density. Hence, actual measurements have shown that the 'hole' formation is *not* an ad hoc assumption of a new process, but was actually observed in the magnetospheres. Other observations in different energy regions show similar phenomena, although sometimes blurred, presumably by radial electric field drifts (Schardt and McDonald, 1983; Krimigis *et al.*, 1983).
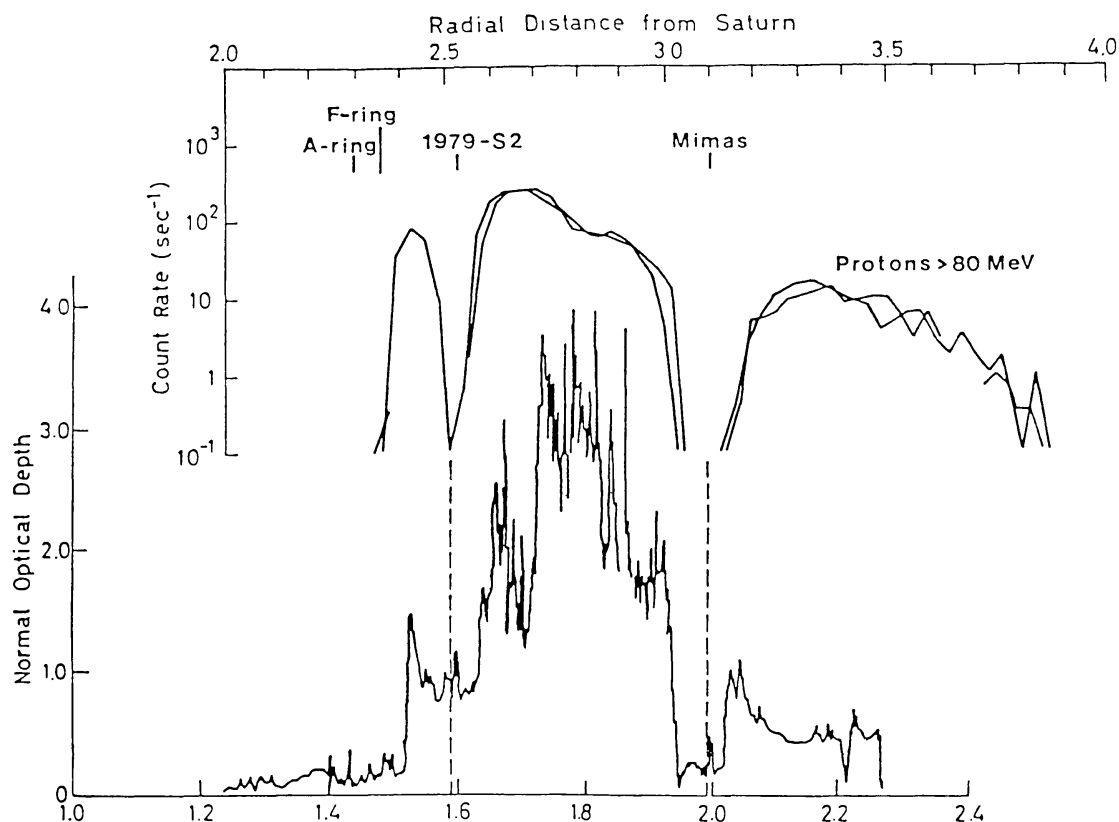

Fig. 8. Comparison between present magnetospheric plasma distribution and mass distribution in the rings. Present-day charged particle distribution in the Saturnian magnetosphere often shows void regions produced by absorption by the satellites. The upper curves are obtained by Fillius and McIlwain (1980). It is argued that the plasma distribution was qualitatively the same in cosmogonic times. The contraction by a factor $\Gamma = 2:3$ at the transition from plasma to planetesimal should result in a somewhat similar mass distribution at a Saturnian distance of $2:3$ of the plasma distribution. The lower curve shows the present mass distribution in the Saturnian rings (Holberg, 1983). It is compared with the Fillius–McIlwain curve reduced by a factor $\Gamma = 0.64$. The 'cosmogonic shadows' of Mimas, the co-orbital satellites and the Shepherds are identified with Cassini's division, the deepest minimum in the B ring, and the inner limit of the B ring.

When, at the plasma-planetesimal transition the 2 : 3 contraction takes place, the hole should be transferred to $\frac{2}{3}$ of the distance of Mimas. This is indeed the location of Cassini's division. Similarly, Janus should produce a marked minimum which now should be found at $\frac{2}{3}$ of their distance. This can be identified with the Holberg minimum* in the B ring (Holberg *et al.*, 1982). The shepherd satellites should produce a hole in a similar way. The A ring also produces a shadow which extends so far out that it joins the shadow produced by the shepherd satellites.

The result of the combined action of the shepherd satellites and the A ring (and perhaps also of the F ring) is an extended region of low intensity which accounts for the C ring. Hence the rapid decrease in intensity at the border between the B and the C ring should be $\frac{2}{3}$ of the position of the shepherds. Finally, the outer edge of the very massive B ring should give the strong decrease intensity which marks the inner edge of the C ring, which is located at $\frac{2}{3}$ of the outer limit of the B ring.

The combined actions of these four cosmogonic shadows give the bulk structure of the Saturnian ring (see Figure 9; for details see Alfvén, 1983b, c).
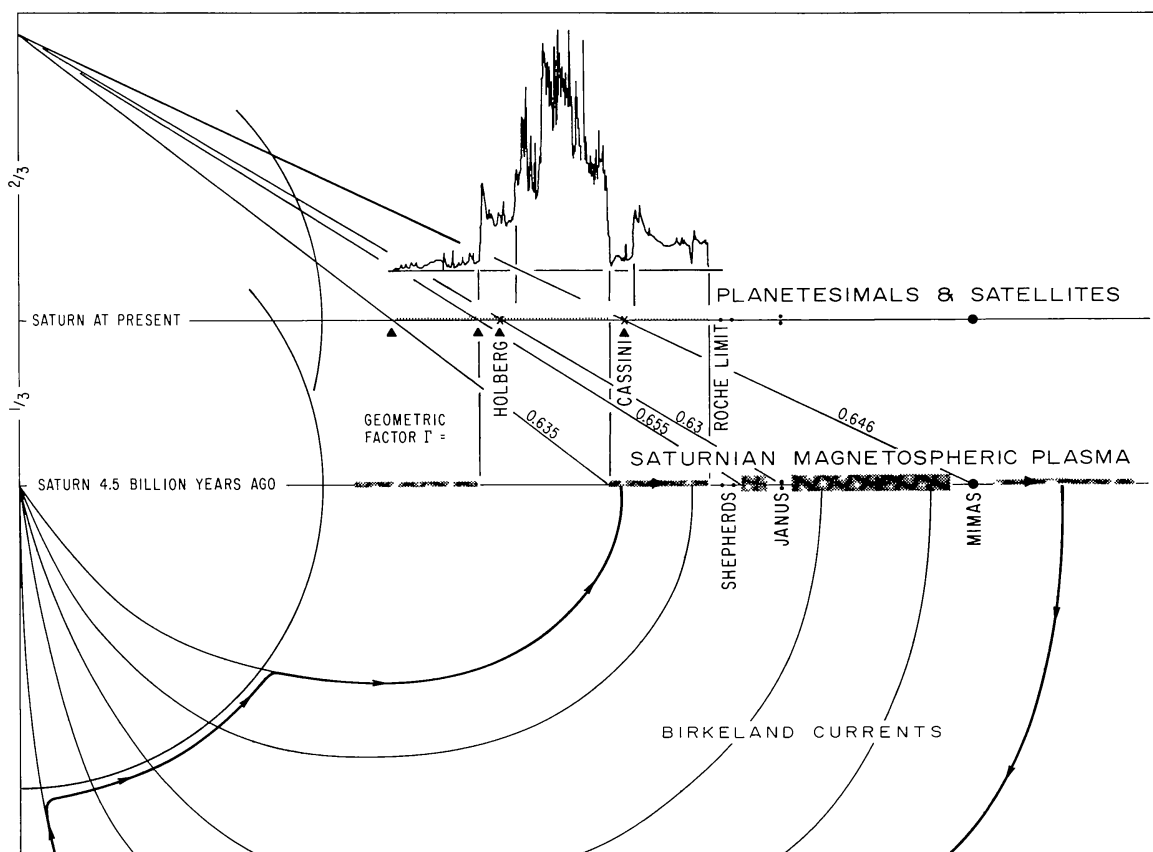

Fig. 9.   Bulk structure of the saturnian ring as a product of the 2 : 3 contraction at the plasma/planetesimal transition. The cosmogonic shadow: – of Mimas produces the Cassini division; – of Janus produces the Holberg minimum; – of the shepherds (and the outer A limit) produces the B–C intensity drop; – of the outer B limit produces the inner limit to C.

* A proposed name for minimum at 1.58 Saturnian radii, motivated by the fact that J. Holberg has drawn attention to its importance in this connection.

When the ring particles produce the shadows, it implies that they absorb plasma with a smaller angular momentum. This leads to a decrease in their distance. Hence, the contraction factor $\Gamma = \frac{2}{3} = 0.667$ should decrease by a few percent. Theoretically we should expect the Γ value to go down to 0.63–0.65 (Alfvén, 1954, 1981b). (In Figure 9 the value 0.64 is used instead of 0.67. This 4% difference is theoretically motivated, but since we never claim a higher accuracy than a few percent, it is not very important for the main discussion.)

3.4. COMPARISON WITH THE ASTEROIDAL BELT

The hetegonic principle states that all systems of secondary bodies should follow a general theory. This means that our results for the Saturnian rings should also be applicable to the asteroid belt around the Sun, which is the other important case where the initial planetesimals have not accreted to larger bodies (see Figure 11; cf. Section 2.5).

A similar analysis of the asteroidal belt has given three identifications of cosmogonic shadows (see Table IVb). The small correction to Γ should not be applied to the asteroidal belt (see Alfvén and Arrhenius, 1976, 11.8, 18.5, 18.8; Alfvén, 1983c).

Also, in this case the mutual agreement and the agreement with theory is surprisingly good.

3.5. ROSSELAND MECHANISM

The basis for the derivation of the $2:3$ contraction has been that electrically charged particles to a first approximation are bound to move along a magnetic field line (seen from a coordinate system in which the electric field perpendicular to B is zero). This is correct as long as their Larmor radii are small compared to the relevant size parameter. However, we have changed the model from the early simple picture of an electron-ion plasma to a treatment of the behaviour of charged dust grains in a dusty plasma. The dust grains have q/m (charge to mass) ratios, which are orders of magnitude smaller than for electrons or protons. In fact, their Larmor periods may exceed the Kepler period and their Larmor radii may be large compared to the characteristic parameters of the dipole field. This seems to make it difficult to treat their motions as we have done. However, this difficulty is often fictitious, as is shown by the following simple model.

In a plasma atmosphere consisting of electrons and ions with mass m_e and m_i, the scale-height

$$h = \frac{2kT}{g(m_e + m_i)} \tag{1}$$

is a compromise between the large-scale height of electrons and the small-scale height of ions. This compromise is produced by an electric field E, called the Rosseland field, or after the Pannekoek-Rosseland field which supports the ions and presses down the electrons. The total burden of the atmosphere is carried by the surface which supports the atmosphere.

We now return to the cosmogonic case:

Consider a cloudlet containing N grains m^{-3} – all of them having the same mass M – in a homogeneous magnetic field B_y. We irradiate the grains with ultraviolet light, so that each emits n electrons, which gives them a positive charge $q = ne$. The electron density is $N_e = Nn$ m^{-3}. This produces a plasma with the Debye distance

$$\lambda_D = (\varepsilon_0 kT_e/Nne^2)^{1/2} . \tag{1}$$

We apply a gravitational force $\mathbf{f}_z = \mathbf{g}M$ on each grain. If λ_D is very large, this would put the grain in motion with the guiding center velocity

$$\mathbf{v}_g = -\mathbf{B} \times M\mathbf{g}/qB^2 \tag{2}$$

in the x-direction, corresponding to a kinetic energy

$$W_g = \tfrac{1}{2}Mv_g^2 = \tfrac{1}{2}M^3 g^2 q^{-2} B^{-2} . \tag{3}$$

When the force is applied slowly, so that the Larmor radius is negligible, the grain is displaced

$$\varDelta z = \frac{W_g}{f_g} = \frac{1}{2} M^2 gq^{-2}B^{-2} . \tag{4}$$

If $\lambda_d \gg \varDelta z$ all the grains will move with velocity v_g, hence producing a current with density

$$\mathbf{i}_g = Nq\mathbf{v}_g . \tag{5}$$

The cloudlet will be prevented from falling by the force per volume from the grains:

$$\mathbf{F} = \mathbf{i}_g \times \mathbf{B} . \tag{6}$$

Because of their small mass, the electrons will not be displaced noticeabley.

This is the same result as if we had only one proton and one electron.

However, the picture is surprisingly different if instead we consider a dusty *plasma*, which is characterized by $\lambda_D \ll \varDelta z$. Then the falling of the grains will produce a positive charge at the bottom and a negative charge at the top of the cloud, producing an electric field E which is so strong that it compensates gravitation.

$$-qE = Mg . \tag{7}$$

This is analogous to the Rosseland field mentioned above.

As the resultant force on the grains is zero, they will remain at rest. The electrons will be put into motion with the velocity $v_e = E/B$, thus producing a current with density

$$\mathbf{i} = -Nq\mathbf{v}_e \tag{8}$$

which is the same as the grain current (according to Equation (5)) would be without the Rosseland field. It is necessary in order to prevent the cloud from falling, because it is not supported by a surface as in the case of an ionized atmosphere.

If grains in an ion-electron plasma get negatively charged by absorbing all – or a considerable part of – the electrons, the same phenomenon takes place, but instead of the electrons the ions will carry the cloud. Hence, if in these idealised cases a gravitational field is applied,

(a) the plasma will be electrically polarized;

(b) the grains remain at rest;

(c) gravitation is compensated by an electron (or ion) current which carries the cloud;

(d) the force is transmitted from the electrons (ions) to the grains by means of a Rosseland field.

The general case of a dusty plasma is of course much more complicated.

Hence a disturbing force will *not* necessarily produce any motion of the grains perpendicular to the magnetic field: the grains are protected from the perturbation by motions of the electrons and ions. The grains can oscillate freely along the magnetic field lines as if they were locked on certain field lines. A small Larmor period or Larmor radius is not necessary for locking them. In a way, the cloudlet behaves as if the grains had no mass. However, a condition is that the current which suspends the cloud is closed in some way.

Application to the 2:3 *contraction mechanism.* A grain which is acted upon by gravitation and centrifugal force can remain 'locked' to a certain magnetic field by means of the mechanism we have studied. The vector sum f_z of the gravitation force and the centrifugal force produces an electric current $\mathbf{i}$ according to (8), which gives a force

$$\mathbf{f} = \mathbf{B} \times \mathbf{i} \tag{9}$$

perpendicular to the field B. For the grain to remain at rest, $\mathbf{f}_z$ must equal f. This means that f_z must be perpendicular to B which defines the size of the centrifugal force, and hence the rotational velocity of the grain.

4. Conclusions about the Plasma-Planetesimal Transition

We have shown that essential features of the bulk structure of the Saturnian rings can be understood as a result of the cosmogonic shadow effect produced by a $\frac{2}{3}$ contraction which probably took place at the transition from the plasma to the planetesimal phase (PPT), presumably 4–5 billion years ago. The 2:3 ratio appears in four cases. Adding to this the three cases from the asteroidal belt, we have no less than *seven identifications*. This means that we can state with considerable confidence that *the cosmogonic shadow effect must have been essential at the formation of the solar system.*

Besides the cosmogonic shadow effect, gravitational resonances were important, especially in the asteroid belt, where they produce the Kirkwood gaps. Due to the very small ratio of Mimas/Saturn in comparison to Jupiter/Sun, corresponding resonance effects in the Saturnian rings are small but clearly identifiable (Holberg *et al.*, 1982). (Also, some other effects are important (see Alfvén, 1983c).) Hence a 2:3 contraction should be characteristic for the plasma-planetesimal transition (PPT). It seems difficult to interpret the above results as being due to any other effect. The surprisingly high
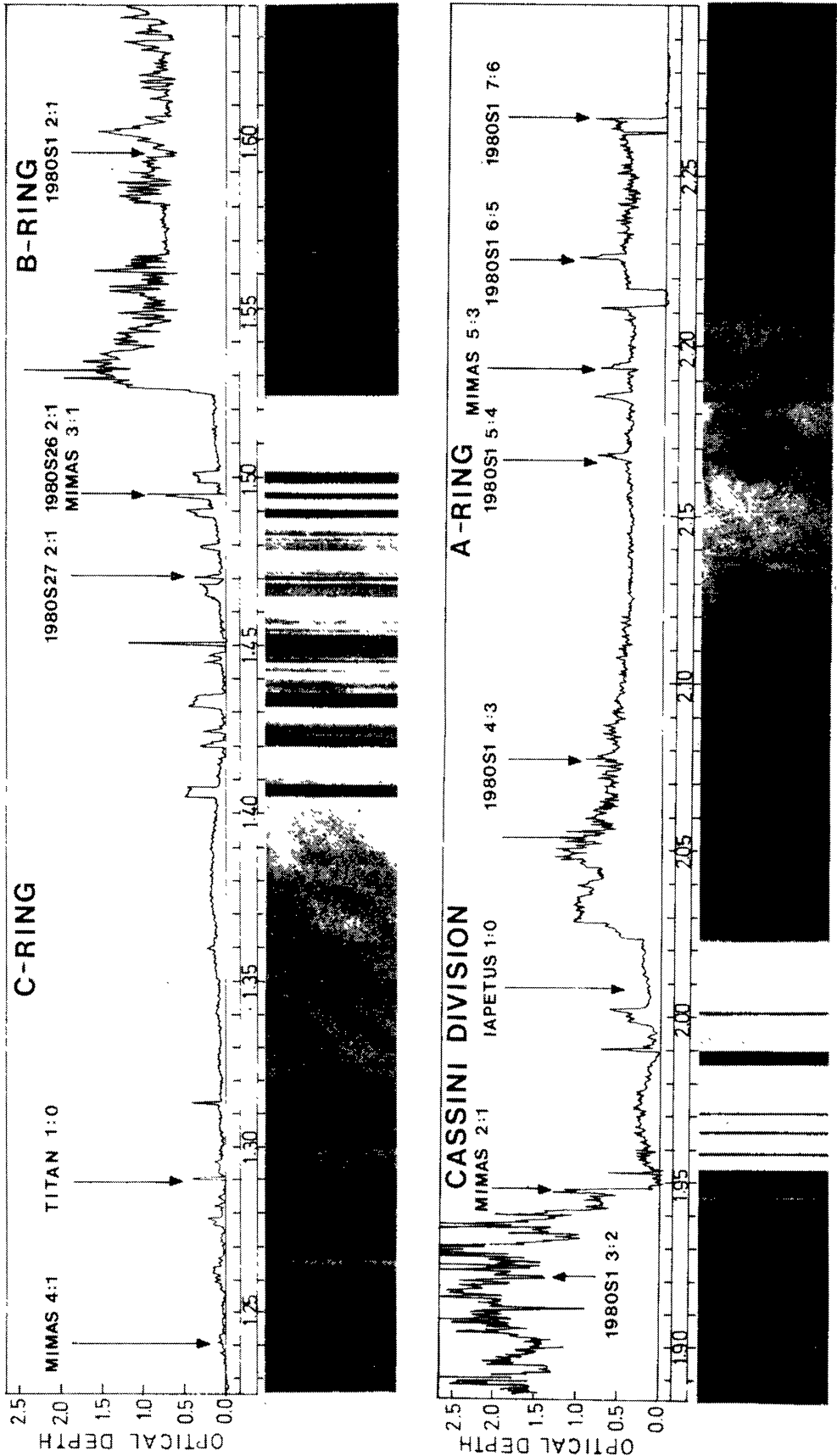
Fig. 10. The normal optical depth of the rings from *Voyager 2* UVS ring occultation data. Below is the brightness of the rings in transparent light (Holberg *et al.*, 1982).
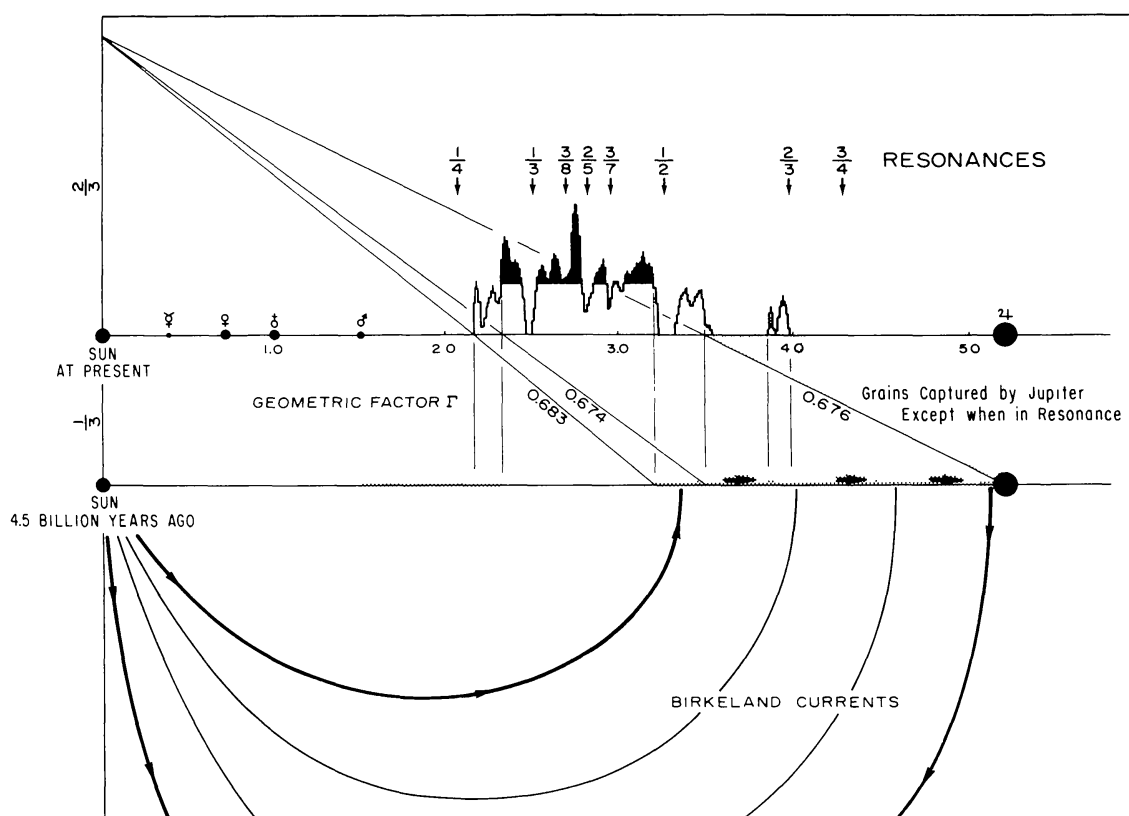
Fig. 11. Asteroid belt. *Diagram of mass distribution* (logarithmic scale). *At the top*: Jupiter's gravitational resonances produce very strong Kirkwood gaps (and the Hilda maximum at $a = 3.9$). More than 98% of the total mass is contained in a *massive section* (darkly shaded area with a sharp inner limit at 2.36 and a sharp outer limit at 3.22. The whole *main belt* has a sharp inner limit at 2.20 and a sharp outer limit at 3.50. *Interpretation*. The shadow of Jupiter ($a = 5.18$) gives the outer limit to the main asteroid belt (3.50). The shadow of this outer limit gives the inner limit to the massive section (2.36). The gravitational resonance 1 : 2 gives the outer limit (3.22) to the massive section, and the shadow of this gives the inner limit (2.20) to the whole main belt (cf. Table IV). The small maximum at $a = 3.9$ represents the Hilda group in resonance with Jupiter.

degree of agreement between the observational and theoretic values means that we have a possibility of *reconstructing certain features of the PPT with an accuracy within a few percent*.

The work on doing this is in progress. A straightforward development of the theoretical model shows that the 'holes' produced by plasma absorption by the satellites must be associated with radial electric fields both at the inside and outside of the depleted region. These electric fields will change the Γ-values so that both the Cassini division and the Holberg minimum should exhibit 'walls' (i.e., maxima) both inside and outside of the minima. Such structures are actually visible in the Holberg curves (Figures 8 and 10).

Further, important features in the bulk structure of the rings (discussed in Alfvén, 1983c) are analyzed, including why the center of the B ring is so massive, why there is a structure inside the Cassini division, and how the several maxima in the C ring are produced. Also, the importance of the Northrop and Hill (1983) instability and the

possibility of explaining the Encke division as produced by an adiabatic circularization of the initial orbits of the grains are being studied. In this way it may be possible to reconstruct the state of the Saturnian environment at cosmogonic times.

With a similar analysis of the asteroidal region, we will have the possibility of reconstructing how the planetesimal state developed into planets/satellites.

Further, we can conclude:

Since the PPT, the structure of the Saturnian rings and the asteroid belt cannot have undergone any violent large-scale disruptions. A slow evolution, resulting in concentration of mass in a few bodies, is indicated. There could have been *no 'solar gale'* strong enough to disrupt the basic pattern. During this time, including the PPT epoch, *there could not have been any strong, large-scale turbulence.*

Violent events before the PPT cannot be excluded, but so far there seems to be no decisive arguments in favor of such phenomena. On the contrary, what has been said in Section 2 speaks against it. The large-scale evolution of the solar system, from an interplanetary cloud to the present structure, could very well have been a slow, quiet process (but consisting of rapid, consecutive processing of cloudlets) which, during many millions of years, built up the present structure.

5. Summary of the Processes Necessary to Understand the Evolution of the Solar System

As was stated in Table II, a cosmogony taking account of electromagnetic effects requires these processes:

(a) Electromagnetic transfer of angular momentum.

(b) The existence of a phenomenon later called 'critical velocity'.

(c) A plasma-planetesimal transition associated with a $2:3$ contraction which produced 'cosmogonic shadows'.

The existence of all three predicted processes has been confirmed. When elaborating the theory in different respects it was found necessary to introduce two more processes:

(d) Local pinch effects in the interstellar source cloud is required to form a non-Laplacean solar nebula of a kind which could give the required initial condition for the cosmogonic processes. This has been discussed in Section 2.1.

(e) A 'Rosseland mechanism' for support of the grains, even if their Larmor radius is very large. The theory of such a mechanism is given in Section 3.5.

Of these, (b), (c), and (e) have been discovered after being predicted from the cosmogonic theory, whereas (a) and (d) have been discovered independently.

So our claim, that the accuracy of our result could not be uncertain by more than a few percent, seems to be legitimate.

The usefulness of a theory is often judged by the number of earlier unknown phenomena it predicts. If we apply this creterion to the introduction of electromagnetic effects in the cosmogonic processes, the electromagnetic approach seems to be not too bad.

Acknowledgments

References

Alfvén, H.: 1942, *Stockholms Observatoriums Annaler I* **14**, No. 2.
Alfvén, H.: 1954, *On the Origin of the Solar System*, Oxford University Press.
Alfvén, H.: 1981a, *Cosmic Plasma*, D. Reidel Publ. Co., Dordrecht, Holland.
Alfvén, H.: 1981b, *Astrophys. Space Sci.* **79**, 491.
Alfvén, H.: 1982, *Physica Scripta* **T2/1**, 10.
Alfvén, H.: 1983a, *Geophys. Res. Letters* **10**, 487.
Alfvén, H.: 1983b, *Astrophys. Space Sci.* **97**, 79.
Alfvén, H.: 1983c, 'Space Research and Cosmic Plasma Physics', lecture at the IUGG Meeting, Hamburg, Aug. 22, 1983, TRITA-EPP-83-06 (Preprint series of the Royal Institute of Technology, Stockholm).
Alfvén, H. and Arrhenius, G.: 1975, *Structure and Evolutionary History of the Solar System*, D. Reidel Publ. Cok, Dordrecht, Holland.
Alfvén, H. and Arrhenius, G.: 1976, *Evolution of the Solar System*, NASA SP-345, US Government Printing Office, Washington, D.C.
Anderson, O. A., Baker, W. R., Bratenahl, A., Further, H. P., Kunkel, W. B., and Stone, J. A.: 1959, *J. Appl. Phys.* **30**, 188.
Axnäs, I., Brenning, N., and Raadu, M. A.: 1982, 'The Critical Ionization Velocity: A Bibliography', in G. Haerendel and E. Möbius (eds.), *Proceedings of the Workshop on Alfvén's Critical Velocity*, Oct. 11–13, 1982, Max-Planck-Institut für Extraterrestrische Physik, Garching, F.D.R.
Baxter, D. and Thompson, W. B.: 1971, 'Jetstream Formation through Inelastic Collisions', in T. Gehrels (ed.), *Physical Studies of Minor Planets*, NASA SP-267, Washington, D.C., U.S. Government Printing Office, p. 319.
Baxter, D. and Thompson, W. B.: 1973, *Astrophys. J.* **183**, 323.
Boström, R.: 1974, in B. M. McCormac (ed.), *Magnetospheric Physics*, D. Reidel Publ. Co., Dordrecht, Holland.
Dessler, A.: 1984, 'Evolution of Arguments Regarding Existence of Field-Aligned Currents', in T. A. Potemra (ed.), *Magnetospheric Currents*, Geophysical Monograph No. 28, American Geophysical Union, Washington, D.C.
Esposito, L. W., O'Callaghan, M., Simmons, K. E., Hord, C. W., West, R. A., Lane, A. L., Pomphrey, R. B., Coffeen, D. L., and Sato, M.: 1983, *J. Geophys. Res.* **88**, 8643.
Fahleson, U. V.: 1961, *Phys. Fluids*, 123.
Fillius, W. and McIlwain, C. E.: 1980, *Geophys. Res.* **85**, 5803.
Haerendel, G.: 1983, *Z. Naturforschung* **37a**, 728.
Hill, T. W., Dessler, A. J., and Goertz, C. K.: 1983, in A. J. Dessler (ed.), *Physics of the Jovian Magnetosphere*, Cambridge University Press, Cambridge, p. 353.

Holberg, J.: 1983, private communication.

Holberg, J., Forrester, W. T., and Lissauer, J. J.: 1982, *Nature* **297**, 115.

Iijima, T. and Potemra, T. A.: 1978, *J. Geophys. Res.* **83**, 599.

Krimigis, S. M., Carbary, J. F., Keath, E. P., Armstrong, T. P., Lanzerotti, L. J., and Gloeckler, G.: 1983, *J. Geophys. Res.* **88**, 8871.

Lin, R., and Bodenheimer, P.: 1981, *Astrophys. J.* **248**, L83.

Potemra, T.: 1979, *Rev. Geophys. Space Phys.* **17**, 640.

Schardt, A. W. and McDonald, F. B.: 1983, *J. Geophys. Res.* **88**, 8923.

Zmuda, A. J. and Armstrong, J. C.: 1974, *J. Geophys. Res.* **79**, 4611.