Q. J. R. astr. Soc. (1993) **34**, 1–20

The Solar System—its Origin and Evolution

M.M. Woolfson
Physics Department, University of York, York YO1 5DD

(Received 1992 July 30; in original form 1991 September 17)

SUMMARY

Scientifically-based cosmogonic theories have been in a constant state of development for the last 250 years. The progress of ideas is reviewed and the problems which arose with early theories are seen to recur in different forms as new theories are proposed. Five current theories are examined: three of these—the Protoplanet Theory, the Modern Laplacian Model and the Capture Theory—are comparatively little known although they offer detailed explanations for the major features of the Solar System. The others are the Accretion Theory and the widely-accepted Solar Nebula Disk Model on which most cosmogonists work at the present time. Finally there is presented a scenario for the evolution of the Solar System which explains most of its important features as a consequence of a collision between early protoplanets.

1 INTRODUCTION

Attempts to explain the origin of the Earth, or of the Solar System, go back to antiquity. For the last 250 years cosmogony has been the subject of intense scientific investigation covering a wide range of possible mechanisms. Here, a number of different models which have been proposed shall be considered, roughly in historical order, but restricted to the basic necessities of a successful theory—the formation of the Sun, the planets and satellite families. In so doing there will be no attempt to cover completely all the work which has been done or to delve into the fine details of individual theories; a full review up to 1985 has been given by Brush (1990) and a Conference Proceedings presented by Weaver & Danly (1989) completely covers work being done on the Solar Nebula Disk Model, the theory which currently engages the attention of most cosmogonists. The goal here will be rather to compare and contrast the general approaches and to highlight their successes and remaining difficulties. Finally a scenario will be presented for the evolution of the Solar System which might follow any suitable formation model and which seems to explain many of the detailed observed features.

2 THEORIES OF THE 18TH CENTURY

In 1749 Buffon suggested that the material that formed planets was knocked out of the Sun by a colliding comet—clearly the exact nature of comets was unknown to Buffon. Laplace (1796) pointed out that any ejected material would move in a Keplerian orbit and eventually would re-enter the Sun although he conceded that the extended nature of the two bodies might negate that conclusion. Anyway, Laplace was convinced that the present near circular orbits of the planets was a necessary condition of their origin.

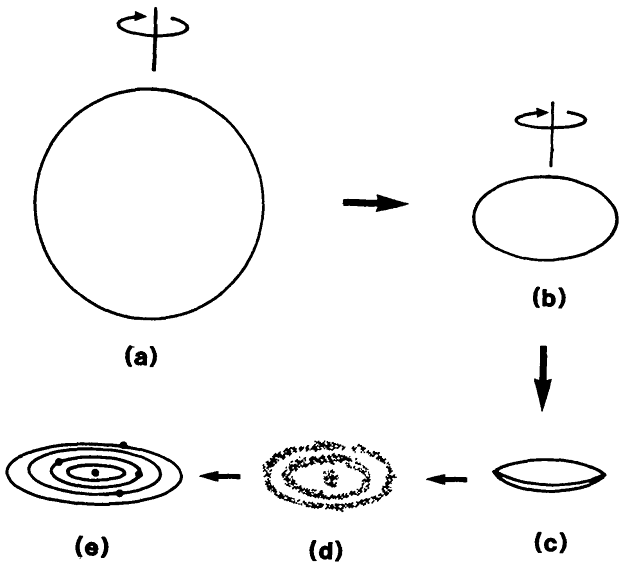

FIG. 1. The Laplace nebula model: (a) a rotating nebula; (b) the collapsing nebula flattened along its rotation axis; (c) formation of a lenticular shape; (d) a series of rings left behind by the contracting core; (e) one residual condensation in each ring forms a planet.

The philosopher Immanuel Kant (1755) theorized that the fuzzy images then being seen in telescopes were either **island universes** (galaxies) or possibly dusty clouds which were in the process of forming stars and attendant planets.

Laplace (1796) gave scientific form to the latter interpretation (Fig. 1). In his model the nebula began as an approximately spherical slowly-rotating dusty cloud. As it cooled and collapsed so it spun faster to conserve angular momentum and flattened along the spin axis. Eventually material at the equator would have been moving in free orbit around the central mass. Further collapse led to material, in the form of annular rings, being left behind in circular Keplerian motion about the interior mass. Condensations within each ring, all orbiting at slightly different rates, gradually accumulated to form a single planet. A smaller scale version of the process for each collapsing protoplanet then yielded satellite systems.

First doubts about the nebula model were based on the distribution of angular momentum in the Solar System since the Sun, with 99·86 per cent of the total mass of the system, contains less than 1 per cent of the total angular momentum, the rest residing in planetary orbits. It seemed impossible for a nebula to evolve in a way that would give this peculiar partitioning of mass and angular momentum. Roche (1854) tried to rescue the Laplace model by postulating a rigidly rotating low-mass nebula enveloping a pre-existing condensed Sun. This highly centrally-condensed starting configuration would have led to a more extreme partitioning of mass and angular momentum but it raised other problems. Jeans (1919) showed that in order for planets to condense from outer nebula material in the presence of the solar tidal field the density had to be more than 0·361 of the mean density of the whole system—implying a large nebula mass.

The conditions for solving the angular momentum distribution problem and for producing planets seemed to be mutually incompatible and so the theory fell out of favour, at least for a while.

3 THEORIES OF THE 20TH CENTURY

The intractable problems of the nebula hypothesis led to a return to two-body theories (as was Buffon's) for which part of the angular momentum problem is avoided by assuming a pre-existing slowly-rotating Sun. Chamberlin (1901) and Moulton (1905) contributed to a model which started with a very active Sun periodically erupting material in the form of huge prominences. Tides due to a passing massive star gave the outcome shown in Fig. 2. Small bodies, called planetesimals, would have formed in the condensations and these eventually aggregated to give planets. The theory was purely descriptive but it provided the foundation for a more substantial theory which followed shortly afterwards.

In 1917 Jeans proposed the tidal model, illustrated in Fig. 3, involving an interaction between the Sun and a massive star. Each condensation in the filament formed a planet which was pulled into elliptical orbit by the gravitational attraction of the star. These initial orbits were rounded off by the effect of a resisting medium consisting of additional material drawn from the Sun. At the first perihelion passage of each protoplanet a small-scale version of the tidal process would have given rise to a satellite family. Jeans supported his model with a considerable body of theory which confirmed both the form of the tidal filament and the gravitational instability which produced condensations within it.

It was Jeffreys (1929), a strong early supporter of the theory, who first expressed doubts. He was concerned about both the likelihood of a very massive star passing close to the Sun and also the spin rates of planets. He invoked a theorem which suggested that since the mean densities of Jupiter and the Sun are similar then they should have similar spin rates instead of being different by a factor of 70 or so. To solve this problem Jeffreys suggested that the star could have struck the Sun a glancing blow, which would have imparted extra spin to the removed material. This was a return to a Buffon-type model.

In due course other difficulties were raised. Russell (1935) showed by a simple argument that a planet produced by tidal disruption could not end up more than a few solar radii from the Sun, a result later confirmed by computer modelling (Lyttleton 1960). The angular momentum problem now appeared in a different guise—not enough was available for planetary orbits.

Spitzer (1939) pointed out that material drawn from the Sun would have a temperature of about 10^6 K and that even for a Jupiter mass its thermal energy would so much exceed the magnitude of its gravitational self-potential energy that it would quickly dissipate into space. It is ironic that the quantitative expression for this balance of energies which threw doubt on the tidal model is due to Jeans—the so-called Jeans criterion.

Jeans (1919) seems to have anticipated these problems because, quite early on, he introduced an amended version of the tidal theory in which the Sun was an extended cool object with a radius equal to that of Neptune's orbit. Russell's and Spitzer's objections no longer apply to this modified model.

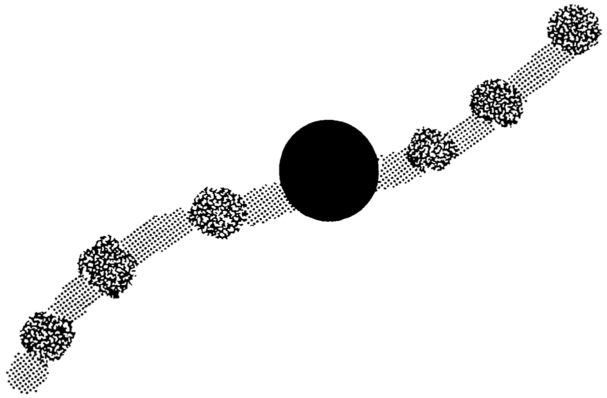

FIG. 2. The Chamberlin and Moulton planetesimal theory. Each high-density region comes from a different solar prominence.

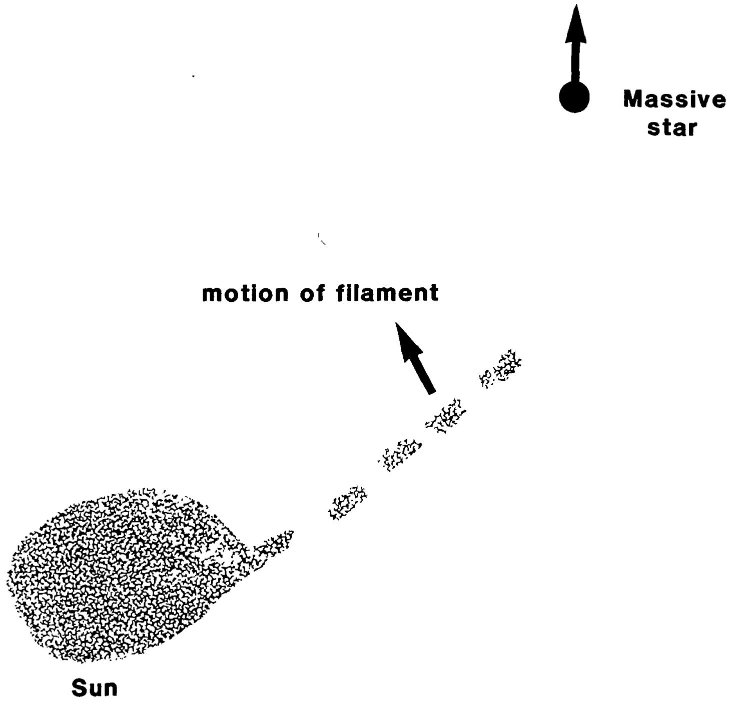

FIG. 3. A representation of the Jeans tidal theory.

Nevertheless difficulties remained, for now the newly-formed planets had to plough through the body of the Sun in their elliptical orbits. Jeans himself accepted that there were problems and wrote "The theory is beset with difficulties and in some respects appears to be definitely unsatisfactory".

4 HIGHLIGHTED PROBLEMS

The angular momentum problem, in some guise or other, emerges as the dominant one in all the early work. By separating Sun and planet formation some theories avoid the difficulty of the slow spin of the Sun but then the problems both of forming planets and of getting them into orbit at the right distances from the Sun seem unresolved. Jeans' extended Sun and cold filament offered the attractive possibility of having both extended orbits and material both dense enough and cool enough to form protoplanets but this promise was confounded by the mechanics of the subsequent motion.

Ingredients for a plausible theory can be deduced from this early work. A slowly rotating Sun must be formed in some way; then dense cool material, preferably in filamentary form so that a limited mass is required, should be placed in orbit around it at some sufficiently great distance. If the protoplanets begin in eccentric orbits then the Jeans tidal mechanism will lead to satellite families and a resisting medium will eventually give near-circular orbits for all the secondary bodies.

While these conditions would give a successful outcome that is not to say that other conditions could not do so. We shall now look at the progress of cosmogonic theories in the last 50 years and see to what extent the difficulties of earlier theories have been resolved. For each of these newer theories there is given in Table I a summary of its successes and failures in solving important problems.

5 THE SCHMIDT–LYTTLETON ACCRETION MODEL

In 1944 the Soviet astronomer Schmidt suggested that the Sun, in much its present form, moved through a dense interstellar cloud and emerged surrounded by a dusty gaseous envelope. Schmidt assumed that a capture process required the presence of a nearby second star to take up energy but Lyttleton (1961) showed that it was not needed, as energy was dissipated in the Sun–cloud interaction. For reasonable postulated properties of the cloud a circumsolar envelope with the mass and angular momentum necessary to form the planets can be accreted. This essentially solves the angular momentum problem—as long as the slow spin of the Sun can be explained in some other way.

The main problem to be solved by this theory is to produce planets. There is a lower limit on the total mass of the envelope based on the need to produce the planets and an upper limit depending on available mechanisms to remove any excess mass. Matter could be lost by evaporation from the outer regions of the envelope, by the action of a strong solar wind during a possible T-Tauri stage of the Sun and by the Poynting–Robertson effect causing fine solid material to spiral inwards. However, there is also a limit on the mass of material which could join the Sun by the last-named process. Just before entering the Sun the material is in free orbit about its equator; since a Jupiter mass ($M_{\odot}/1000$) in orbit around the solar equator would have three times as much angular momentum as the Sun has in its spin then, clearly, the Sun can only have had limited addition to its substance in this way. These considerations give an upper limit for the envelope mass of a few per cent of that of the Sun. Any model which required a more massive envelope would need to explain how it was removed.

The first step in the development of the envelope would be that solid grains settled into the mean plane, as defined by the net spin angular momentum. This disk of grains would have been gravitationally unstable and would have

TABLE I

Successes and difficulties of five modern cosmogonic theories in solving some problems associated with explaining the origin of the solar system

	Accretion Theory	Protoplanet Theory	Capture Theory	Modern Laplacian Theory	Solar Nebula Theory
Slow solar spin	Not considered. Needs auxiliary theory for formation of Sun.	Well explained by process of stellar formation.	Explained by star formation model.	Moment-of-inertia factor reduced by needle-like prominences. A.M. transported outwards.	Unexplained
Distribution of mass and angular momentum	Well explained by characteristics of the captured cloud.	Explained by theory as given. Some doubt about availability of A.M. for planetary orbits.	Explained by filament and Sun–star orbit characteristics.	Explained by shedding of concentric gas rings.	Several suggested mechanisms for A.M. transfer but none completely satisfactory.
Formation of terrestrial planets	May accrete on time-scale 10^6–10^7 yr.	Retained portions of fissioned protoplanets.	Collision of protoplanets in terrestrial region.	Short timescale for accretion in a ring.	May accrete on timescale 10^6–10^7 yr.
Formation of major planets	Unacceptable timescales for outer planets.	Well explained. Retained major portion of fissioned protoplanets.	Condensations in a dense filament.	Short timescale for accretion in a ring. (apparent problem with ring stability).	Unacceptable timescales for outer planets. For massive disk, mass-disposal problem.
Formation of satellites	Not considered.	Droplets between fissioned protoplanets.	Sun-protoplanet tidal interaction.	Small-scale version of planet formation. Good predictions of orbits and compositions.	Not considered in detail.
Explanation of Bode's law	None.	None.	None.	Well-explained.	None.

clumped together to form planetesimals of dimensions of 1 km, more or less. The way in which the planetesimals would then combine to give planets was first investigated by Safronov (1972). He showed that mutual perturbations would lead to eccentric orbits and hence collisions between planetesimals. A collision between two bodies would lead to fragmentation of one or both of them and then to dispersal or aggregation of their material. Aggregation is more likely if one of the colliding bodies is much more massive than the other. In any region of the disk one body would tend to become dominant and so accrete all other mass in its region. Timescales calculated by Safronov for forming terrestrial planets were about 10^7 yr but to produce the cores of major planets took much longer—as much as 10^{10}–10^{11} yr for Neptune—at least twice the age of the Solar System.

While the Accretion Theory provided at least a partial solution to the angular momentum problem it also introduced the difficult problem of producing planets from initially diffuse material.

6 THE PROTOPLANET THEORY

McCrea (1960) put forward the Protoplanet Theory which, as a central feature, explained both the slow rotation of the Sun and the formation of planets. The model begins with a dense interstellar cloud which is going to form a stellar cluster. As it collapsed so it became turbulent and colliding streams of turbulent material created dense regions which moved haphazardly in the less dense background material. These were termed floccules; when they collided they coalesced and about 20 of them would have formed a stable aggregate according to Jeans' criterion. Here and there in the cloud an aggregate would have formed of sufficient mass to act as a substantial gravitational attracter and this would eventually have become a star. Smaller aggregates would then have been captured in orbit around the star to form a planetary system.

There are many attractive features of this model. Because of the way a star would form, by the accretion of floccules coming from random directions, it would have comparatively little angular momentum, a few times that of the Sun perhaps. This is not a serious discrepancy since there are very plausible ways in which an early Sun could have lost a modest amount of angular momentum. One way is where ionized solar wind material spirals around magnetic field lines and corotates with the Sun out to a few solar radii. By the time the material decouples from the field it has transported angular momentum out of the Sun. A reasonable enhancement over present values of both the solar wind and magnetic field in an early active Sun could easily permit an initial angular momentum complement some ten times that of the present Sun.

A prediction of the model is that the planets would have possessed too much spin angular momentum. McCrea has turned this apparent drawback into an asset. Using an idea due to Lyttleton (1960) McCrea describes how a collapsing major planet with sufficient angular momentum would eventually break up into two parts with a mass ratio greater than 8:1. The smaller part would have had a high speed relative to the centre of mass of the combined bodies and would have left the system. The remaining larger fragment would be left with much less angular momentum. Fluid droplets formed between the two fragments as they separated would have been retained by the larger fragment to form a satellite system. A similar fission process in the inner part of the Solar System, where escape velocities are higher, would not have led to the loss of the smaller fragment. McCrea suggested that the pairs Earth–Mars and Venus–Mercury arose in this way and noted that the densities of the combined pairs of bodies would be similar.

McCrea was able to get good numerical agreement between various features of the solar system and the model predictions. However, the original theory was criticized on the grounds that the floccules would have a very short lifetime ($\sim$ 1 yr) and would not have lasted long enough to form stable aggregations. McCrea (1988) has overcome this by modifying the model so that the original high density regions are stable and of protoplanetary mass, formed by fragmentation of the cloud rather than by supersonic collisions of turbulent gas streams.

The mechanisms in the theory have not been subjected to detailed modelling and there is at least one feature which can only be examined in this way. In assessing the amount of angular momentum which would reside in the planetary orbits McCrea assumed that the cloud consisted of isolated regions within which angular momentum was conserved. Thus any angular momentum in a region which did not end up in the spin of the star was considered as residing in the orbits of a planetary family. However, protoplanets would have moved between the regions and, indeed, may have joined a forming star in a neighbouring region. This gives the possibility that some, or even most, of the angular momentum not contained in the spin of the star might have ended up in the relative motion of stars—and hence would not have been available for planetary orbits.

Here we do, at last, have a theory which explains the basic characteristics of the Solar System although many questions still need to be addressed. There is a need to study the mechanisms operating in a collapsing cloud. Will protoplanetary condensations be the first to form or would they be of greater, even stellar, mass? How likely is it that the planets would all orbit in the same direction and that the solar spin axis would be at only 7° to the normal to the mean plane of the system? Would there be enough angular momentum for the planetary orbits or would most of it appear in the relative motion of stars? The Protoplanet Theory is susceptible to numerical modelling, which would answer the above questions and either put the theory on a sounder footing or, perhaps, throw doubt on its plausibility.

7 THE CAPTURE THEORY

The Capture Theory (Woolfson 1964; Dormand & Woolfson 1989) involves a tidal interaction between a condensed star (Sun) and a collapsing low-density protostar. Although, as now developed, it begins with the interstellar medium and ends with the solar system in much its present form the present description will be limited to the formation of the Sun, planets and satellites.

The process begins with a supernova which injected heavy elements into, and hence cooled, a region of the interstellar medium. Iodine/Xenon dating from meteorites suggest a time of somewhat less than 2×10^8 yr between that event and the eventual formation of the solar system. Material was drawn into the low-pressure, cool region until, finally, a cool dense cloud was formed in pressure equilibrium with the surrounding less-dense but hotter interstellar medium and in thermal equilibrium with galactic radiation. The mass of the cloud was above the Jeans critical mass and so it collapsed, slowly at first but at an accelerating rate. Turbulence was generated in the cloud, fed by the release of gravitational energy, and collisions between turbulent elements produced regions of higher density. Re-expansion was a slower process than cooling so that cool dense regions resulted. At a certain stage in the overall cloud collapse a dense region would itself have collapsed on a timescale less than that for the remixing of its material back into the cloud. At this stage the first stars were formed. The result of the total process was the formation of a galactic stellar cluster. The numerical modelling of this process (Woolfson 1979) gives results agreeing with the following observations:

(1) The sequence of formation of stars according to mass as observed in young galactic clusters.
(2) The mass index, i.e. number density of stars as a function of mass.
(3) The angular momentum of stars over the total mass range. This includes slow spin for late-type stars like the Sun.
(4) Maser observations of star forming regions giving the dimensions of the total star-forming region, the size of individual sources (a forming star) and turbulent speeds.

A consequence of this mode of star formation is that interactions between stars become possible, or even probable. The basic capture-theory process, illustrated in Fig. 4. involves an interaction between a compact star of solar mass and a diffuse star of lesser mass. This model does however conflict with the McCrea protoplanet scenario; long before protoplanetary condensations were formed, if ever they did, there would have been stars around.

The protoplanets produced in the capture theory model, estimated as six in number, would have collapsed on short timescales of a few tens of years. The Sun would have induced rotating tidal bulges on them which during the final stages of collapse, would have become distorted into a filament-like form. Condensations in these filaments gave regular satellite families and the characteristics of the satellite families of Jupiter, Saturn and Uranus are quantitatively well explained by such an origin (Williams & Woolfson 1983).

Many theorists have concluded that the relationship of planets to the Sun and satellites to their parent planets are so similar that the two kinds of system should have come about by the same basic process. Indeed, Galileo was confirmed in his belief in the Copernican model by his first view of the Galilean satellites of Jupiter. Jeans (1919) stated "...; any hypothesis which assigned different origins to the main system and the sub-systems would be condemned by its own artificiality" and Alfvén (1978) similarly records his view "We should not try to make a theory of the origin of planets around the Sun but a **general theory of the formation of secondary bodies around a central body**". For an earlier source of such thoughts we must go back to Newton's Principia where he gives Rules of Reasoning in Philosophy, the second of which is "**Therefore to the same natural effects we must, as far as possible, assign the same causes**". Newton then goes on to give pairs of

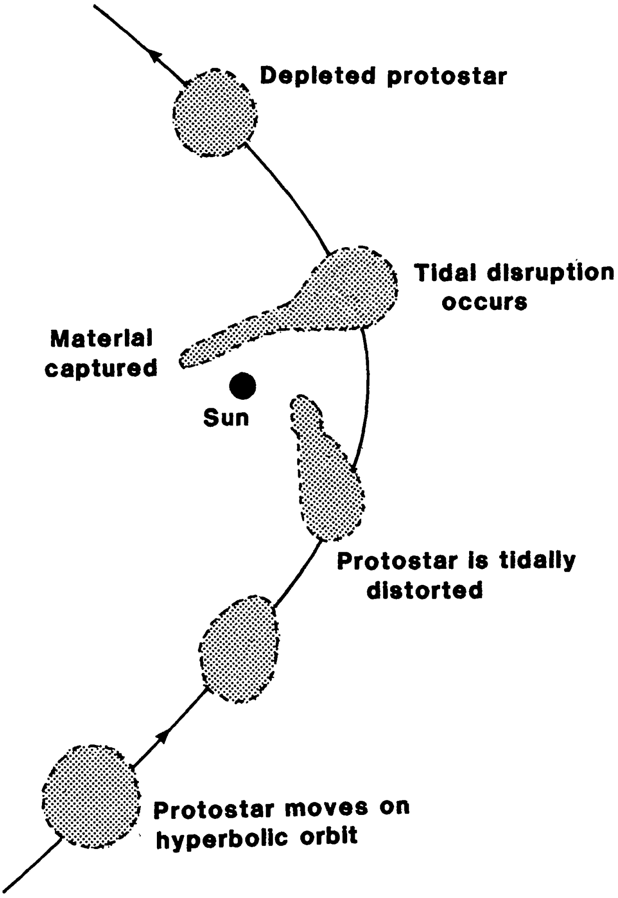

FIG. 4. A schematic representation of the Capture Theory showing a sequence of configurations of the disrupted protostar.

observations to which one should ascribe the same causes one of which is "the light of our culinary fire and of the Sun". We now know that the causes of these two observations are quite different but it does illustrate that one must be careful in applying general principles. It is instructive to look at the planetary and satellite systems in terms of the ratio of the intrinsic angular momentum (i.e. angular momentum per unit mass) of material at the equator of the primary body and of a typical secondary body in its orbit. For the Sun and Jupiter the ratio is 1:7800; for Jupiter and Io it is 1:8. For the Sun and Saturn it is 1:10600; for Saturn and Titan it is 1:11. The planet:satellite ratio, which is similar for each system, is one which can be accommodated comfortably by a common origin of planetary spin and satellite orbit as the capture theory model (and the Jeans tidal theory) suggests. The Sun:planet ratio is very different and it is not necessarily a bad feature of a theory that it should suggest a different origin—indeed, it may even be a good feature.

The elliptical orbits of the protoplanets rounded off in a resisting medium, with an estimated total mass five times that of Jupiter (5 M_J), which formed around the Sun; rounding times varied from $\sim 10^5$ yr for Jupiter to 2×10^6 yr for outer planets. The solid part of the medium, with mass about $M_J/10$, was absorbed by the Sun, due to the Poynting–Robertson effect, so pulling the solar spin axis towards the normal to the mean plane of the system, which accounts for the small but significant angle of 7° between these two directions.

8 A MODERN LAPLACIAN MODEL

A new form of the Laplacian hypothesis has been extensively modelled by Prentice (1978). He takes as an initial requirement that a cloud collapsing towards becoming a star must become rotationally unstable when its radius is approximately that of Neptune's orbit. Normal interstellar dense clouds, under normal collapsing conditions, would have about 100 times too much angular momentum so Prentice calls on a protostar formation model suggested by Reddish & Wickramasinghe (1969). This considers the segregation of solid grains in the cloud, including H, at a very low temperature. Gas drag removed angular momentum as the grains settled; eventually the released gravitational energy vaporized the grains but there was a stellar mass condensation left with one hundredth of the angular momentum given by straightforward gaseous collapse. The subsequent collapse was non-homologous and eventually a state was reached with a dense luminous central core with radius about $3R_{\odot}$ and a tenuous protostellar envelope of density 10^{-10} kg m^{-3}.

As previously noted, Roche had attempted to rescue the Laplace model by postulating a very centrally condensed starting point. The protostellar configuration we are now considering goes some way towards this goal but not far enough. At this stage Prentice (1978) introduces the concept of supersonic turbulent stress which he associates with a very active T-Tauri state of the early Sun. Needle-like supersonic convective elements were ejected through the photosphere at speeds between 100 and 200 km s^{-1} and the material returned by descending more slowly at 10 km s^{-1}. Prentice (1978) notes that this process would explain the observations of outflow from T-Tauri stars but without the need for mass loss. The effect of the process was a dramatic lowering of the moment-of-inertia factor, which is what is required to explain the partitioning of angular momentum.

At the outer limit of the convective-overshoot region a dense shell of quiescent material would have formed. As the protostar collapsed and the angular speed increased so this shell concentrated more and more in the equatorial zone until, at some stage most of the mass in the shell resided in a torus or ring (Fig. 5). Eventually the configuration became rotationally unstable, the torus separated and a new torus began to form. Prentice (1978) has shown that the spacing of the rings gives an explanation of Bode's law, an empirical law which describes the progression of the radii of the planetary orbits, which also applies in a modified form to the spacings of satellite systems.

The temperatures in the rings depended on their radii and were 26 K for the Neptunian ring, 122 K for Jupiter and 1260 K for Mercury. Solid materials condensed out of the rings and the expected chemical compositions of planets seem well explained by this process. The condensates formed in the ring were acted upon by gas drag and settled close to the axis of the ring. This process would have taken a period somewhere between 3000 and 300000 yr for the Jupiter ring. If the linear density of this axial material was high enough then it would have been gravitationally unstable and clumped

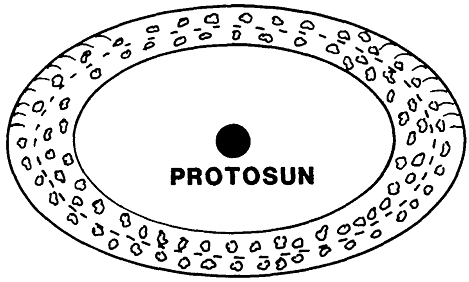

FIG. 5. A gaseous torus around the protosun. Solid grains settle towards the axis (----).

together to form larger planetesimals and finally planets; the timescale for this process was 2000 yr for Mercury, 7000 yr for Jupiter and 90000 yr for Neptune. The major planets would have accreted extensive atmospheres and the collapse of these is presumed to have followed a similar pattern to that described for the total nebula to give satellite families.

Prentice has used his model to make predictions about the detailed composition and chemical structure of solar-system bodies and also about satellites previously unobserved. Thus, before Voyager reached Neptune he had predicted a satellite system which corresponded quite well with those discovered (Prentice 1989). One possible objection to the model is in the lifetime of the rings. Prentice states that their expected lifetimes were about 3×10^5 yr, the duration of the active T-Tauri stage of the Sun. However, it is difficult to see how the detached Neptunian ring, for example, was sustained by the prolonged solar activity once it became detached. The minor radius of the Jupiter torus was less than 0·2 AU and an approximate calculation, similar to that which gives Jeans' criterion, suggests that at a temperature of 122 K a torus density of about 10^{-5} kg m^{-3} would be required to avoid rapid expansion. The density given by Prentice, 10^{-7} kg m^{-3}, seems far too low to have retained the Jupiter torus for even the minimum predicted time of 3000 yr required for condensates to settle to the axis.

The Modern Lapacian Theory has been analysed in considerable detail by Prentice and, of all the cosmogonic models which have been proposed, it is the only one which has been used to make successful predictions.

9 NEW NEBULA IDEAS

Long before Prentice's complete reformulation of the Laplace model there were frequent returns to it with various proposals put forward to solve its difficulties. Berlage (1930) looked at some complicated mechanisms for the evolution of a nebula disk involving viscosity—although this work paid no attention to the basic angular momentum distribution problem. In 1944 von Weizsäcker put forward a model for a disk involving a symmetrical pattern of vortices but again the main thrust of the idea was to give a mechanism for planet formation rather than to solve the angular momentum problem.

In the early 1970s a new intense wave of solar nebula work began and it is now the dominant activity in cosmogony. Part of the original impetus for this activity seems to have come from the community of meteoriticists. The chondrules in meteorites are spherical liquid silicate drops which cooled and became incorporated in a matrix of silicate fragments. In some cavities in carbonaceous chondrites mineral crystals are found that must have condensed directly from the vapour phase. In addition the minerals found in meteorites are well explained as successive condensations in a dense hot nebula.

By 1978 Cameron was putting forward a version of the nebula theory which involved a very massive disk with mass equal to that of the Sun, with planets forming by direct condensation as giant protoplanets with up to 30 times the mass of Jupiter. These large bodies were then assumed to have been broken up by collisions and subsequently the debris collected together again to form a few giant planets and a large number of smaller bodies, the asteroids. This process would have required the disposal of a considerable mass of material but Cameron did not deal with this problem. One of the features of this model is that material falling onto the disk as it is forming gives a great deal of turbulence and hence energy dissipation. Cameron then called on a theoretical result from Lynden-Bell & Pringle (1974) that if a rotating disk evolves in such a way that its energy of motion decreases while its angular momentum remains constant then this is achieved by material close to the spin axis moving inwards while material further out moves outwards. This is tantamount to an outwards transmission of angular momentum. Curiously, in view of what has been indicated previously, a feature of this model is that it does not give the meteoriticists what they want—a hot nebula. Cameron pointed out quite specifically that "At no time, anywhere in the solar nebula, anywhere outwards from the formation of Mercury, is the temperature in the unperturbed solar nebula ever high enough to evaporate completely the solid materials contained in interstellar grains".

The new nebula theorists have produced a large number of papers over the past twenty years or so and it is difficult to cover the complete range of ideas that have been put forward. They point out that the Solar Nebula Disk Model (SNDM) has now acquired an observational basis in the infrared observations of some pre-main sequence stars, which show an infrared excess. This indicates the presence of a low temperature emitter of much larger area than the star itself and a very plausible explanation of this is the presence of a dusty disk. One star, β Pictoris, gives an infrared image which could be interpreted as an oblique view of a disk. The evidence is suggestive but not universally accepted. Nevertheless, one thing that everyone agrees about is that these infrared sources are shortlived and are restricted to pre-main sequence stars of age less than about 3×10^6 yr.

The problem of the transfer of angular momentum, which led to the downfall of the original Laplace model, has attracted a great deal of attention. It should be noted that it is not enough just to extract angular momentum while material is moving inwards to join the central mass, which eventually forms the Sun. Material spiralling inwards is in free orbit around the interior mass; if one makes the untenable but generous assumption that, as the Sun assembles, its density as a function of distance from its centre follows the present pattern then a simple calculation shows that the Sun would end up with several hundred times its present angular momentum. It is thus necessary either to have material moving more or less radially inwards to join the Sun or a mechanism which can remove almost all the angular momentum *after* material has joined the Sun.

There are four basic mechanisms which have been proposed for angular momentum transfer:

1. (1) Turbulent viscosity
2. (2) Gravitational torque
3. (3) Magnetic braking
4. (4) Wave propagation and shock dissipation.

Mechanisms (1), (2) and (4) are of purely mechanical origin and only operate to move some material towards the central condensation while other material is pushed outwards. We have already mentioned the Lynden-Bell & Pringle (1974) turbulent viscosity mechanism. Gravitational torque acts if the collapsing nebula develops trailing spiral arms. The non-axisymmetric potential which arises from the gravitational effect of the spiral arms produces a torque on inner material which reduces its angular momentum and allows the inner material to collapse. Conversely outer material acquires this angular momentum and moves outwards. Based on this idea Miyama (1989) has proposed a complicated scenario for the formation of the Sun and the solar nebula. His model is too complex to analyse and needs to be computed numerically but the process is highly non-linear and cannot be reliably handled with current computers. Against this, Larson (1989) states that this mechanism is ineffective if the mass of the disk is less than about one tenth that of the central star so that some other mechanism is also needed.

The final mechanical process—wave propagation—relies on the fact that waves transport energy and momentum and that in a rotating system they will also transport angular momentum. Thus if acoustic waves are generated in the central part of the system, by turbulence or perhaps supersonic accretion flows onto Jupiter, and these waves propagate without serious energy loss to the outer parts of the system then angular momentum will be transferred outwards.

Magnetic braking can operate on the central core itself, for example by the process previously described where solar wind material corotates with the magnetic field. Another process involves magnetic linkage between the central body and a surrounding disk where the magnetic field is frozen into ionized material at each end (Hoyle 1960); as the central core contracts and tries to spin more quickly so the field lines become stretched, pulling backwards on the core to slow its rotation while providing the disk with more angular momentum so that it moves outwards. However, any reasonable selection of parameters for such a model can only reduce the angular momentum from about ten times the present value to what is now observed—which is what the McCrea model requires.

The problem with forming giant gaseous protoplanets directly from the disk, as suggested by Cameron, is that it requires an extremely massive disk, most of the substance of which must eventually be removed, which requires a considerable source of energy. The alternative theory, that of building upwards from grains, is being intensively worked on at present. It is worth noting that if infrared observations are taken as a basis for supporting the SNDM then they also impose a tight time restraint on planetary formation.

The terrestrial planets must be formed in a few million years, certainly less than 10^7 yr, as must the solid cores of the giant planets, which must then capture gaseous envelopes. Work by Mizuno *et al.* (1978) and by Mizuno (1980) suggested that these cores would all be of similar mass—about ten times that of the Earth but the Jupiter core would take 10^8 yr to form. All theoretical work in planet building suggests that there seems to be only marginal difficulty with timescales in the terrestrial region, which is what Safronov found. The timescale problem for outer planets, which was revealed by Safronov's work, has been tackled by Wetherill (1989) who has described a 'runaway process' whereby a core collects new material to itself at an accelerating rate. While this shortens the timescale considerably the present form of the idea requires a disk of areal density much greater than theory allows.

The timescale problem which has just been described relates to the accumulation of planetesimals into terrestrial planets or giant-planet cores. There must also be considered the prior stage in going from grains to planetesimals. The process of forming the latter by gravitational instabilities has been subjected to some criticism by Weidenschilling *et al.* (1989). They state that the dust layer would be extremely sensitive to turbulence which would prevent gravitational instability effects unless coagulation of grains forms bodies large enough ($\sim$ 1 m) to decouple from the gas. In fact Weidenschilling and his coworkers think that grain adhesion processes can obviate the need for gravitational instability altogether.

Wetherill (1989) considers that the problem with explaining the building of planets is the complexity of the physical and chemical processes involved in the necessary steps in the sequence of planet growth. Workers on the SNDM tend to concentrate on a limited aspect of the total problem so the assumptions and conclusions of each piece of work may not necessarily be consistent with those of all other workers. By the investigation of ranges of possible mechanisms for all stages of the SNDM, including the distribution of angular momentum, there may eventually emerge some mutually consistent components of a total theory starting with the collapsing nebula and finishing with the Sun, planets and satellites.

10 A MODEL FOR THE EVOLUTION OF THE SOLAR SYSTEM

In 1971 Dormand & Woolfson concluded that Capture Theory planets could not form in the terrestrial region as the protoplanets would have approached the Sun too closely and been torn apart. Later, in 1977, when they were numerically investigating the round off of a system of non-coplanar protoplanets they found an effect which resolved the difficulty of the terrestrial planets—and, as it turned out, many other problems as well. Due to the mass of the resisting medium the net force on the planets was not centrally directed which caused a precession of the orbits. Thus the early non-coplanar protoplanet orbits, which intersected in projection and precessed at different rates, would have intersected in space from time to time giving the possibility of major interactions. The characteristic times for major planetary interactions were found to be similar to round-off times so that interactions were probable. Tidal effects, through close passes, can explain the various directions of planetary spin axes, in particular that of Uranus which is almost in the plane of its orbit and slightly retrograde. Another postulated interaction is a direct collision between two large planets which, had they survived, would have rounded off in the regions of the asteroid belt and Mars.

Numerical models of collisions show that a range of conditions can be found for which the more massive of the two planets (A) would have been expelled from the solar system while the less massive one (B) was sheared into two almost-equal parts, the non-volatile portions of which rounded off in the orbits of the Earth and Venus. The colliding planets were those closest to the Sun and each would have had several large satellites. Many possible scenarios exist where satellites can become associated with one or other of the fragments; for example, a satellite of planet A can be tightly captured by the Earth fragment in an orbit of eccentricity 0·4. This idea fits with the Moon's observed hemispherical asymmetry, corresponding to a thinning of the crust on the hemisphere facing the earth. Satellites of the colliding planets would have suffered severe abrasion of the hemispheres which pointed towards the collision. Connell & Woolfson (1983) have proposed that Mars, also showing hemispherical asymmetry, was a large satellite which was retained in heliocentric orbit and they have also explained the distribution of its surface features in relation to its spin axis. Mercury could also have been another ex-satellite, with its crust and mantle so heavily abraded that, when it re-assembled in spherical form, its density was similar to that of the Earth.

Dormand & Woolfson (1980) have also modelled the situation where Triton, an escaped satellite, interacted closely with Pluto a prograde satellite of Neptune. Pluto was ejected into an orbit of similar characteristics to its present one, while Triton was captured in retrograde orbit about Neptune. Pluto's orbit would have precessed while the resisting medium was present so that it no longer intercepts that of Neptune.

A planetary collision would have given a large range of temperature conditions which would explain many physical characteristics of iron and chondritic meteorites, including carbonaceous chondrites. Achondrites may then be identified as unmetamorphosed crustal material situated at some distance from the collision region. It is tempting also to think of icy surface material, spalled off the remote faces of the planets, as a potential source of an inner cometary cloud, perturbation of which by stars, giant molecular clouds and the galactic tidal field can feed and maintain the Oort cometary cloud.

Some samples of meteorites show interesting isotopic anomalies. Oxygen is found which resembles terrestrial oxygen with the addition of pure ^{16}O, while excess ^{26}Mg in some samples can certainly be related to the previous presence of ^{26}Al, which has a half-life of 720000 yr. It has been suggested that a supernova injected this into the solar system immediately prior to its birth. However, the anomaly which imposes the tightest constraint is so-called Neon E which is highly enriched, or even pure, ^{22}Ne. The most probable source is radioactive ^{22}Na with a half life of 2·6 yr. Since the sodium would have to be contained in a mineral in a cool rock within a few years of its formation a supernova origin immediately before the formation of the solar system seems unlikely. Clayton (1975a, b) has suggested that isotopic anomalies were present in interstellar grains which became incorporated into solar system material long after the nucleosynthetic events which produced them. Another possibility we examine here is that they were actually formed in the solar system.

Planets formed by condensations in a cold tidal filament would have contained a complete cosmic mix of elements. Jupiter would have retained virtually all the initial material since even hydrogen would not have enough thermal energy to escape and the D/H ratio in Jupiter will be the same as that in the cosmos at large, estimated as 2×10^{-5}. The D/H ratio on Earth is some eight times this value, in some meteorite samples it is six or more times the Earth value and on Venus it is one hundred times the Earth value, or $1{\cdot}6 \times 10^{-2}$. The reason for the enrichment of deuterium relative to hydrogen is fairly straightforward; an atmosphere containing hydrogen compounds, for example water, would be dissociated by solar radiation and hydrogen atoms may escape. However if it is HDO which dissociates then the deuterium atom, with twice the mass of hydrogen can be retained. The evolution of planetary atmospheres has been studied over wide ranges of conditions—planetary mass, thermal regime etc.—and moderate mass planets over a wide range of conditions are found to lose most of their atmospheric hydrogen to give highly enriched D/H ratios (Michael 1991). When the planets cool this high deuterium ratio may exist within icy materials incorporated into silicate surface layers.

Calculations, admittedly at this stage somewhat crude, suggest that in the contact region of a planetary collision a temperature in the range $2\text{–}5 \times 10^6$ K will be produced. This temperature suffices to trigger deuterium reactions and thence a whole chain of nuclear reactions involving all the types of atom available. Solving the equations for these reaction chains, involving 90 isotopes up to sulphur plus iron (which plays no part in the reactions but acts as a coolant) involved in 138 nuclear reactions, and allowing for all factors such as ionization, shows that a final temperature of 5×10^8 K can result together with the production of ^{16}O enrichment, ^{26}Al and ^{22}Na. Of course there are other products as well and at the present time the relationship of these to possible observations has not been investigated. The expanding vaporized material from the collision region expands and cools on a timescale of a year or less. The formation of condensed and cool silicate materials containing sodium minerals, including ^{22}Na, on a timescale of a few years seems quite feasible in this scenario.

This evolutionary hypothesis, which depends on the high original eccentricity of protoplanetary orbits, a resisting medium to give precession and a consequent collision, is associated with the Capture Theory model but can be appended to any other theory which gives similar initial conditions.

11 CONCLUSIONS

The last 30 yr have seen intense activity in the field of cosmogony with the emergence of several new theories and the refurbishment of older ones. This period has also seen the development of the computer as an investigative tool so that even quite complex mechanisms can be subjected to numerical investigation—an advantage which the pioneers in the field did not enjoy. It is no longer necessary to make approximations to find analytical solutions to equations describing proposed processes; more often than not they can be confirmed as either plausible or implausible.

The angular momentum problem has been tackled successfully by the Protoplanet Theory, the Capture Theory and the Modern Laplacian Theory. The SNDM has not yet come forward with a convincing solution to the angular momentum problem and there are grounds for preferring the Accretion Theory, since plausible mechanisms for producing a slowly rotating Sun by independent means do exist and the two theories have in common the problem of obtaining planets from diffuse material.

The SNDM can easily explain the production of giant protoplanets, as Cameron originally suggested, but there are unsolved problems with the grains–planetesimals–planet scenario. There is no compromise state for the disk—it must be massive enough to produce planets directly by gravitational instability, which is a well-founded process, but of low enough mass to be able to dispose of excess material. The Protoplanet and Capture theories both produce planetary-size bodies from cool and dense material by gravitational instability in a way which gives no disposal problem. The Modern Laplacian theory also offers a scenario for producing planets although it is not clear that the conditions to produce them will be of sufficient duration.

A valid question to ask of any theory is that concerning the predicted frequency of planetary systems which would result. This begs the question of what constitutes a planetary system—for example would a stable three-body system with a main-sequence star and two brown dwarfs be one? It seems very likely that if there are many planetary systems then they will have a wide range of parameters. The Protoplanet Theory, the SNDM and the Modern Laplacian Theory, which all link directly the processes of star and planet formation, would seem to predict that systems were common. The Capture Theory is less direct in its linkage, although all the events happen within a stellar cluster, but if a wide variety of systems exist in terms of scale and composition then it too will make them reasonably common—associated with, perhaps, 1 per cent or so of stars.

The possible existence of planets associated with the neutron star PSR 1257+12 (Wolszczan & Frail 1992) is taken as evidence that it must be possible to produce planets from diffuse material since this is all that would remain in the aftermath of a supernova. The logic of this is uncertain; a system with major planets at distances of several hundred astronomical units from the parent star might well give surviving non-volatile planetary cores which could then spiral inwards under the influence of the supernova debris. This is something which could and should be subjected to modelling.

It is in the nature of the cosmogonic problem, even more than for science in general, that the term 'correct' cannot be applied to any theory. The only judgements that can be made are of relative plausibility, with some ideas so implausible that the term 'incorrect' becomes acceptable. For this particular area of science Georges Lemaitre (1950) has written "The purpose of any cosmogonic theory is to seek out **ideally simple conditions** which could have initiated the world and from which, **by the play of recognized forces**, that world, in all its complexity, may be resulted". The bold type is not original but emphasizes what seem to be the important parts of Lemaitre's statement.

The science that is used should be well-understood and readily acceptable; if resort must be made to complicated and convoluted arguments or special phenomena invented solely for the purpose of explaining some aspect of the theory, then confidence in it should be low.

There is no reason to suppose that the next 30 yr will not be as interesting and productive as those which have just passed; new ideas will arise and new judgements will be made.

REFERENCES

Alfvén, H. 1978. *The Origin of the Solar System*, p. 41, ed. Dermott, S.F., Wiley, Chichester.

Berlage, H.P. 1930. *Proc. K. Ned. Akad. Wet.* **33**, 614.

Brush, S.G. 1990. *Rev. Mod. Phys.* **62**, 43.

Buffon, G.-L. Leclerc, Comte de, 1749. *Histoire Naturelle Générale et Particulière avec la Description du Cabinet du Roi*, Imprimerie Royal, Paris.

Cameron, A.G.W. 1978. *The Origin of the Solar System*, p. 49, ed. Dermott, S.F., Wiley, Chichester.

Chamberlin, T.C. 1901. *Astrophys. J.* **14**, 17.

Clayton, D.D. 1975a. *Astrophys. J.* **199**, 765.

Clayton, D.D. 1975b. *Nature*, **257**, 36.

Connell, A.J. & Woolfson, M.M. 1983. *Mon. Not. R. astr. Soc.* **204**, 1221.

Dormand, J.R. & Woolfson, M.M. 1971. *Mon. Not. R. astr. Soc.* **151**, 307.

Dormand, J.R. & Woolfson, M.M. 1977. *Mon. Not. R. astr. Soc.* **180**, 243.

Dormand, J.R. & Woolfson, M.M. 1980. *Mon. Not. R. astr. Soc.* **193**, 171.

Dormand, J.R. & Woolfson, M.M. 1989. *The Origin of the Solar System: the capture theory*, Ellis Horwood, Chichester.

Hoyle, F. 1960. *Q. J. R. astr. Soc.* **1**, 28

Jeans, J.H. 1917. *Mem. R. astr. Soc.* **77**, 186.

Jeans, J.H. 1919. *Problems of Cosmogony and Stellar Dynamics*, Cambridge University Press, Cambridge.

Jeffreys, H. 1929. *Mon. Not. R. astr. Soc.* **89**, 636.

Kant, I. 1755. *Algemeine Naturgeschichte und Theories des Himmel*, (Königsberg 1975); *Universal natural history and theory of the heavens*, English translation and introduction by Jaki, S.L. (Edinburgh 1981).

Laplace, P.S. de, 1796. *Exposition du Systéme du Monde*, Imprimerie Cercle-Social, Paris.

Larson, R.B. 1989. In *The Formation and Evolution of Planetary Systems*, p. 31, eds. Weaver, H.A. & Danly, L., Cambridge University Press, Cambridge.

Lemaitre, G.E. 1950. *Primeval Atom: an Essay on Cosmogony*, Van Nostrand, London.

Lynden-Bell, D. & Pringle, J.E. 1974. *Mon. Not. R. astr. Soc.* **168**, 603.

Lyttleton, R.A. 1960. *Mon. Not. R. astr. Soc.* **121**, 551.

Lyttleton, R.A. 1961. *Mon. Not. R. astr. Soc.* **122**, 399.

McCrea, W.H. 1960. *Proc. R. Soc. Lond.* **A250**, 245.

McCrea, W.H. 1988. *The Physics of the Planets*, p. 421, ed. Runcorn, S.K., Wiley, Chichester.

Michael, D. 1991. *D.Phil. Thesis*, University of York.

Miyama, S.M. 1989. In *The Formation and Evolution of Planetary Systems*, p. 284, eds. Weaver, H.A. & Danly, L., Cambridge University Press, Cambridge.

Mizuno, H., Nakazawa, N. & Hayashi, C. 1978. *Prog. Theor. Phys.* **60**, 699.

Mizuno, H. 1980. *Prog. Theor. Phys.*, **64**, 544.

Moulton, F.R. 1905. *Astrophys. J.* **22**, 165.

Prentice, A.J.R. 1978. *The Origin of the Solar System*, p. 111, ed. Dermott, S.F., Wiley, Chichester.

Prentice, A.J.R. 1989. *Phys. Lett.* **A140**, 265.

Reddish, V.C. & Wickramasinghe, N.C., 1969. *Mon. Not. R. astr. Soc.* **143**, 189.

Roche, R. 1854. *Mem. Acad. Montpellier*, **2**, 399.

Russell, H.N. 1935. *The Solar System and its Origin*, Macmillan, New York.

Safronov, V.S. 1972. *Evolution of the Protoplanetary Cloud and Formation of the Earth and Planets*, Israel Program for Scientific Translations, Jerusalem.

Schmidt, O.Y. 1944. *Dokl. Akad. Nauk USSR*, **45**, No.6.

Spitzer, L. 1939. *Astrophys. J.* **90**, 675.
von Weizsäcker, C.F. 1944. *Z. Astrophys.* **22**, 319.
Weaver, H.A. & Danly, L., (eds) 1989. *The Formation and Evolution of Planetary Systems.* Cambridge University Press, Cambridge.
Weidenschilling, S.J., Donn, B. & Meakin, P. 1989. In *The Formation and Evolution of Planetary Systems*, p. 131, eds. Weaver, H.A. & Danly, L., Cambridge University Press, Cambridge.
Wetherill, G.W. 1989. In *The Formation and Evolution of Planetary Systems*, p. 1, eds. Weaver, H.A. & Danly, L., Cambridge University Press, Cambridge.
Williams, S. & Woolfson, M.M. 1983. *Mon. Not. R. astr. Soc.* **204**, 853.
Wolszczan, A. & Frail, D.A. 1992. *Nature*, **355**, 145.
Woolfson, M.M. 1964. *Proc. R. Soc. Lond.* **A282**, 485.
Woolfson, M.M. 1979. *Phil. Trans. R. Soc. Lond.* **A291**, 219.