Bull. Astr. Soc. India (1998) **26**, 339-348

The origin of the solar system*

K.D. Abhyankar

Astronomy Dept., Osmania University, Hyderabad 500007, India

Abstract. This is a historical review of the developments of the ideas about the origin of the solar system starting from the nebular hypothesis of Kant and Laplace. It deals with the tidal hypothesis of Moulton and Chamberlin, the capture hypothesis of Schmidt and the theories of the formation of the planets due to Weizsacker, Kuiper and Hoyle. The relation of the formation of the planetary system to the process of formation of stars is discussed with particular reference to the transfer of angular momentum from the star to the surrounding disc. The review ends with the recent results concerning the discovery of planets and circumstellar discs around young stars.

1. Introduction

The subject of the origin of the Solar System i.e. cosmogony has attracted the attention of man from the most ancient times. *Rigveda* I, 164, 1-6 speaks of the seven luminaries in the sky, viz., the sun, the moon, and the seven planets, as seven horses yoked to the never wasting wheel of time. The phrase "The sages have woven seven threads together to form a web" aptly describes the complicated motion of the seven luminaries on the ecliptic. The rishis also asked the question like: "Who has seen the birth of the first born (the sun)?" and "Who has established and fixed the six regions (above the earth)?" The Vedic literature often refers to the seven regions including the earth. They can be identified as: (1) *Bhooh,* the earth, (2) *Bhuvah,* the intermediate space containing lightening (atmosphere), the moon and Mercury which is called the son of the moon, (3) *Swah,* the heaven containing the Sun and Venus which is called the son of the sun, (4) *Mahah,* the region of Mars which is called the son of the earth, (5) *Janah,* the region of Jupiter, (6) *Tapah,* the region of Saturn, and (7) *Satyam,* the region of the stars.

The Vedic poets could not go beyond expressing their ignorance and curiosity about these matters, because they did not have the benefit of telescopes, backend instruments, CCD's, space platforms, space probes and other modern techniques. However, we can now talk a little more authoritatively on the subject of cosmogony.

---

* Based on Chapter 10 of the author's forthcoming book : *Astrophysics of the Solar System*

2. Common characteristics of the solar system

We shall start by listing the common properties of the solar system. They are:

(1) Common west to east motions of revolution and rotation of all bodies almost in one plane.
(2) Sun has 99.86 per cent of mass, but only one per cent of the angular momentum.
(3) Ages of all the bodies in the solar system are close to 4.5 billion years.
(4) Deuterium (D) and Lithium (Li) in the planets have the interstellar abundance while the sun contains one tenth of that amount.
(5) The inner terrestrial planets are smaller, less massive and more dense while the outer Jovian planets are larger, more massive and less dense.
(6) Titius-Bodes law holds approximately.
(7) Asteroids are found between Mars and Jupiter while comets are situated in the Oort's cloud at distances of 10^5 AU.
(8) Most planets have small satellites and the larger ones have rings around them.
(9) There exist twins such as Mercury-Mars, Earth-Venus, Jupiter-Saturn and Uranus-Neptune.

Any good theory about the origin of the solar system would have to explain these features.

3. The nebular hypothesis

The common west to east orbital and rotational motion of the bodies in the solar system indicates that they have a common source of angular momentum. Kant (1775) suggested that they could have been formed in a rotating gaseous nebula and he conjectured that Andromeda nebula might represent a planetary system in the making. It was thought that the central nucleus would become a sun-like star and the knots in the spiral arms would end up as planets. But we now know that the Andromeda nebula is a large stellar system like the Milky Way containing billions of stars. However, the concept that the planetary system originated in a rotating gaseous mass is accepted and it is known as the nebular hypothesis.

The physics of the contraction of a rotating sphere of gas was considered by Laplace (1798) and Jeans (1919). The angular momentum A of the sphere is $A=I\omega=\alpha MR^2\omega$, where I is the moment of inertia, M the mass, R the radius and ω the angular velocity of the sphere. The parameter α depends upon the central condensation of the sphere. As A remains constant during contraction ω increases as $(1/R^2)$. Consequently the centrifugal force at the equator $F_c=R\omega^2$increases as (I/R^3) while the centripetal gravitation force $F_g=GM/R^2$ increases as $(1/R^2)$. So at a critical value of $R=R_c=A^2/\alpha^2GM^3$ the centrifugal force F_c becomes larger than F_g and a ring of material would be ejected at the equator as shown in Fig. 1 for a Roche model of the star. Further contraction will give rise to ejection of more such rings and each ring would condense into a planet.

There are several objections to the nebular hypothesis as follows: (i) the central body will have the maximum angular momentum which is not at all true for the solar system. (ii) Even if all the angular momentum of the planets is put in the sun it will not be rotationally unstable until it attains a radius of 0.27 $R_\odot$ on contraction. (iii) The shape of the sun would be elongated which it is not.

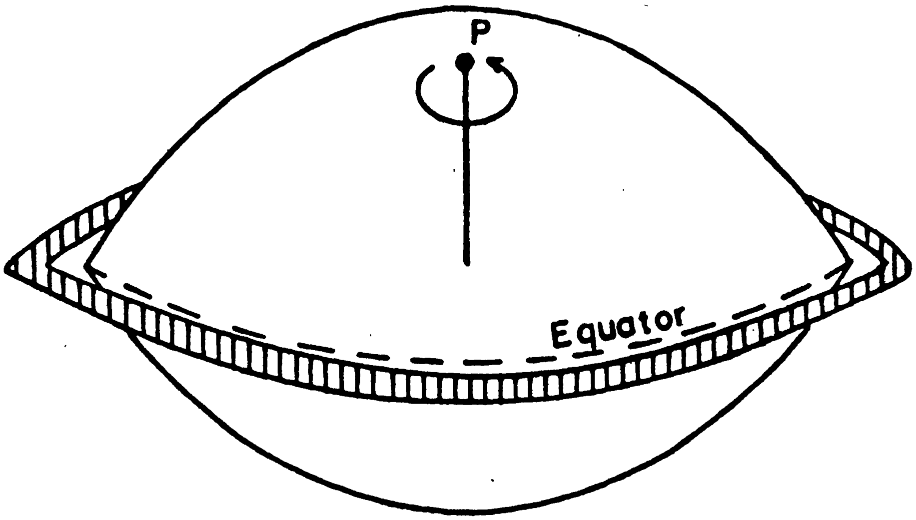

Figure 1. Rotational ejection of a ring in a Roche model.

4. The tidal hypothesis

In order to get over the problem of the angular momentum, Moulton (1905) and Chamberlin (1916) put forward their tidal hypothesis. In their theory it is assumed that a passing star raised a huge tide over the sun and the planets were formed in the tidal filament after the departure of the passing star as shown in Fig. 2. The process of the formation of the tide was considered by Jeans (1929) who concluded that if the sun had uniform density there would be only a few pieces in the tide which would have masses comparable to that of the sun. However, a centrally condensed sun could give rise to small planets. But further dynamical treatment by Jeffreys (1924) showed that the two stars will have to undergo a grazing collision for imparting enough angular momentum to the planets. And Russell (1935) later found that the tidal filament would have to be dragged to a distance of hundreds of astronomical units for the same result. So he

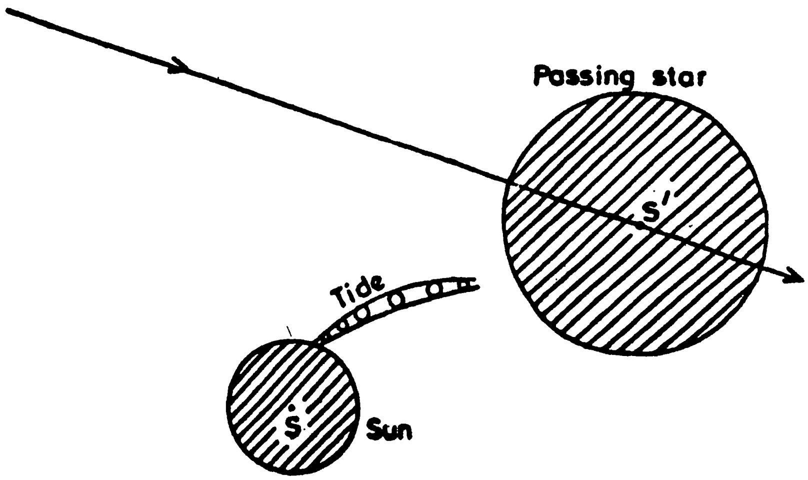

Figure 2. The tidal hypothesis.

suggested that the sun had a companion and the passing star collided with the companion and broke it up. As the planets were formed in that debris they got their angular momentum from that of the companion.

The tidal hypothesis and its modification by Russell have also a few difficulties: (i) Nölke (1930) and Spitzer (1939) concluded that the tidal material would be too hot and it will disperse instead of condensing into planets. (ii) Chandrasekhar (1941) proved that stellar collisions are extremely rare due to the vast distances between the stars of the Milky way. (iii) As D and Li have interstellar abundance while they are depleted in the sun by thermonuclear reactions the material of the planets would never have been a part of the sun. (iv) Urey (1957) gave evidence that the earth started as a relatively cold body and it heated up later by the decay of radioactive elements U, Th and K.

5. The capture hypothesis

In order to take care of the difficulties with the tidal hypothesis Schmidt (1944) suggested that there was a collision between the sun and an interstellar cloud. The sun captured a portion of the cloud and the planets were formed therein. This process provided the required mass and angular momentum to the planets. The capture hypothesis has also the following drawbacks: (i) Star-cloud collision are more frequent than the star-star encounters, but they are also quite rare. (ii) The relative velocity of collision would be of the order of 20 km/s which makes the collision too fast for the capture of the cloud. (iii) The ages of the planets are almost identical with that of the sun, but the capture would make them unequal.

However, Schmidt was able to give the following satisfactory scenario for the formation of planets in the solar nebula. In the beginning the dust would be separated from the gas by the solar radiation pressure and it would form a disc around the sun. Adhesion of the dust particles would give rise to larger particles leading to the formation of protoplanets of the size of the moon. The protoplanets would grow by accretion of smaller fragments and their orbits would become circular by averaging of the eccentricities of the combining bodies. The inner planets would loose the volatiles due to solar heating while the outer planets would be able to retain hydorogen, helium and other lighter gases. That the planets were formed by accretion is now evident from the space studies of planets. The moon and the planets and even satellites are found to have craters produced by the accreted fragments. However, due to the difficulties with the capture hypothesis the problem of the angular momentum remained unsolved.

6. Theories of Weizsacker and Kuiper

Weizsacker and Kuiper skirted around the problem of the angular momentum by assuming that the sun somehow acquired a nebula with adequate mass and angular momentum and they proceeded to consider the formation of the planets in the nebula.

Weizsacker (1945) argued that there will be turbulence in the solar nebula which would give rise to the formation of eddies having angular motion opposite to that of the rotation of the nebula. In order to get Bode's law for which the ratio of the major axes of successive planetary orbits is $a_{n+1}/a_n \approx 1.8$, he conceived a pattern of five eddies in a circle of 360°. He put smaller and smaller eddies in circles of smaller and smaller radii with ball bearing eddies at the boundaries as shown in Fig. 3. The protoplanets were supposed to have formed at the sites of the ball bearing eddies and the merger of the protoplanets in the same orbit produced the known planets.

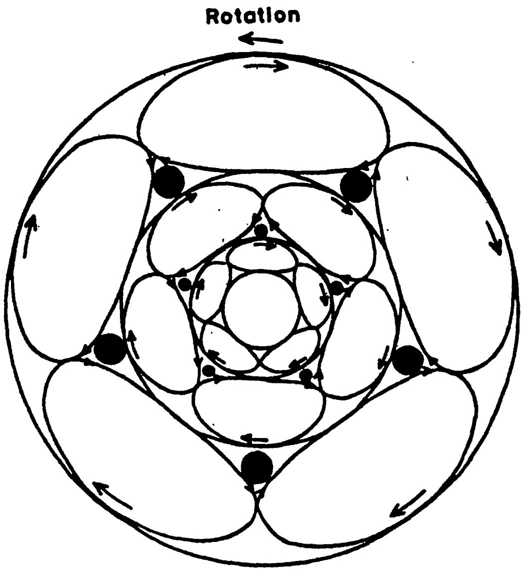

Figure 3. Weizsacker's pattern of eddies.

Weizsacker's theory was challenged by Chandrasekhar (1946) who showed that there will be a wide spectrum of turbulence with smaller eddies within the larger ones and there would be no regular pattern as suggested by Weizsacker. Further the life time of the eddies will be too short for the formation of the planets.

Kuiper (1951) assumed that the planets were formed in their present positions because the density at these distances exceeded the Roche density $\rho_c=16(R_\odot/a)^3\rho_\odot$, where $R_\odot$ is the radius of the sun, $\rho_\odot$ its mean density and 'a' the distance from the sun. Further, he showed that the thickness h of the nebular disc will increase linearly with distance as h=0.0176 AU. Then assuming that the n-th planet was formed in a taurus with outer and inner radii equal to $(a_{n+1}-a_n)/2$ and $(a_n-a_{n-1})/2$ respectively, he calculated the masses of the protoplanets. This gave a total mass of the nebular disc to be of the order of 0.09 to 0.18 solar masses. Obviously 99 percent of the material was lost, more of it being lost by the inner terrestrial planets and less by the outer Jovian planets. This adhoc theory was only able to give the estimate of the mass of the nebular disc and nothing else.

7. Relation to the formation of stars

From a statistical study of the spectroscopic and visual binaries Kuiper (1951) found that their separations ranged between 3×10^{-3} to 10^5 AU with the majority of them having separations between 3 and 30 AU. As this is comparable to the size of the solar system it appears that the formation of the solar system is perhaps a special case of the formation of binaries.

Fig. 4 shows the relation between the total angular momentum A and the total mass M for binaries and single stars. In both cases A is approximately proportional to $M^{5/3}$, but the angular momentum is about 16 times larger for the binaries with respect to the single stars. Further, in the case of single stars the rotational velocity and the angular momentum drops drastically below $3M_\odot$ reaching a very low value for stars like the sun. However, if we add the angular momentum of the planets to that of the sun we get the point S which falls on the linear curve for the massive single stars. This indicates that the smaller stars are perhaps all accompanied by planets on account of some common mechanism. This was identified by Hoyle (1960) as the magnetic coupling between the rotating central star and the ejected nebular disc.

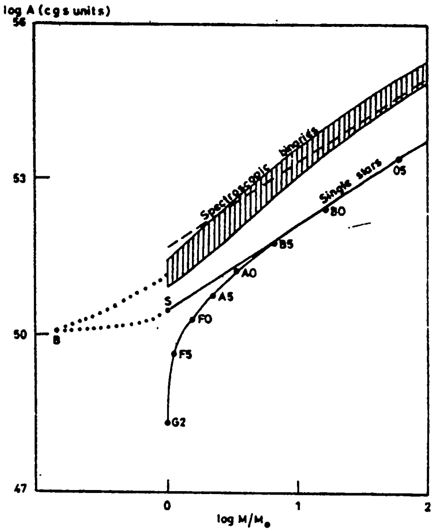

Figure 4. Angular momentum of the binaries and single stars (S=solar system, B = Bernard star system).

8. Hoyle's theory

Stars form when a molecular cloud contracts under its own weight. During contraction it gets fragmented into smaller bits that condense into individual stars. The cloud possesses angular momentum on account of the differential rotation of the galaxy. Consider a cloud of mass M and mean density $\rho = n\mu m_H$ where n is the particle density, μ the molecular weight and m_H the mass of the hydrogen atom. It will have a radius $R=(3m/4\pi n\mu m_H)^{1/3}$. Further, it will have an angular velocity $\omega = A_0 - B_0$ where A_0 and B_0 are Oort's constants of galactic rotation. Putting in the values of n, A_0 and B_0 one finds that $A = 4\times10^{54}(M/M_\odot)^{5/3}$ for a giant molecular cloud. Comparing this with $A=5\times10^{51}(M/M_\odot)^{5/3}$ for binaries we see that 99.9 per cent of the angular momentum of the cloud is lost during fragmentation, most of it going into the relative motions of the individual stars. Anyway, we can start with a fragment with angular momentum appropriate for a binary and we see what happens when it tries to form a single star instead of a binary.

First we note that the critical radius for ejection of a ring occurs at $R_c=4.7\times10^{12}(M/M_\odot)^{1/3}$, which is equal to 0.31 AU for 1 $M_\odot$ and 0.67 AU for 10 $M_\odot$. So a ring is formed and it grows into a disc containing about 10 percent mass of the star. During contraction the magnetic field, which has a value of 10^{-5} G in the interstellar space, gets magnified to 1 to 10 G. Now $T_e \approx 5000$°K during the Hayashi phase of contraction, so the disc is partially ionised. As the ions move along the lines of force as pointed out by Alfven (1954) the disc gets magnetically connected with the central star. As the star contracts further the lines of force are wound up in the convective layers of the star as shown in Fig. 5. The resulting torque transfers about 90 to 95 per cent of the angular momentum of the star to the disc which is pushed out to 10 AU. This process of magnetic breaking of the star is discussed in detail by Schatzman (1962).

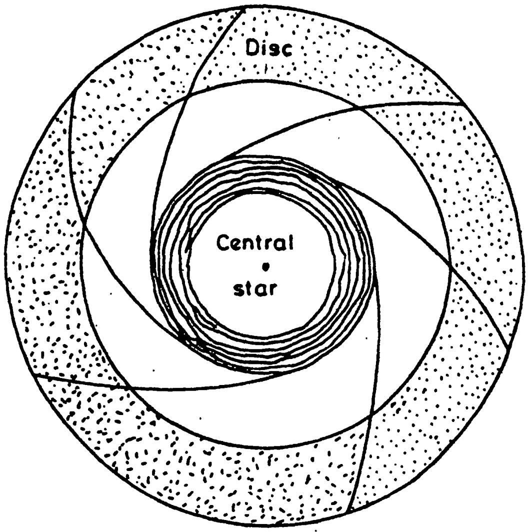

Figure 5. Magnetic coupling between the central star and the disc.

Now in the case of a massive star ($M > 3 M_{\odot}$) the surface temperature quickly rises to values much above 10000°K which makes the star's envelope radiative. The disappearance of the convective envelope stops the transfer of angular momentum to the disc. Further a strong stellar wind disperses the disc into interstellar space. Thus we end up in a rapidly rotating single star.

In the case of less massive stars ($M < 3M_{\odot}$) the depth of the convective envelope increases with the decrease in the mass of the star. Consequently, there is greater transfer of angular momentum to the disc with a decrease in the rotational velocity of the star towards less massive stars. Further, the wound up magnetic field becomes unstable producing flare activity which disperses most of the disc into space. In the case of the sun only one per cent of the original disc is left for the formation of planets.

9. Formation of the planets in the disc

One can trace the following steps in the formation of planets within the disc.

(1) The nebular disc around the sun will have a temperature ranging from 2000°K near the sun to 8°K at 10 AU. This will favour the formation of solid grains of metallic oxides of Fe,

Mg, Si, etc. The grains will be covered by the ices of H_2O, CO_2, CH_4, NH_3, etc. at distances greater than 4 AU.

(2) The grains will grow by adhesion to sizes of 1m to 1cm in 10^5 years. They will further aggregate into larger objects like 100 to 200 km size asteroids and lunar size planetesimals in 10^6 years.

(3) The planetesimals will grow by accretion of smaller fragments to full size terrestrial planets in the inner region in 10^7 years. They will not be able to collect volatile gases due to the higher temperatures in the region of their formation.

(4) In the outer regions where the temperature is low the planetisimals can collect gases to form the Jovian planets. In their case the H/He ratio will be smaller for the less massive planets Uranus and Neptune.

(5) The asteroidal belt between Mars and Jupiter indicates that the growth of planets was halted outside Mars due to the formation of Jupiter. Alternatively an incipient planet might have broken up due to the tidal action of Jupiter at its close approach to that planet. Bombardment of the inner planets by the asteroids in elliptical orbits would have provided them with the volatile gases like H_2O and CO_2.

(6) According to Hughes (1993) not much material was left to form planets beyond Neptune. So the icy bodies which were not incorporated in the Jovian planets were thrown out to form the Oort's cloud of comets.

(7) The satellite systems were formed around the planets by the same process of ejection of rings which formed the planets around the sun. But there was no transfer of angular momentum from the planet to the satellites. Rings were formed by the breaking of the satellites which came within the Roche limit of the parent planet.

The following peculiarities may however be noted:

(i) The ratio of mass of the moon to that of the earth is quite large (1/81) compared to the satellite-planet mass ratio for other satellites. Hence one has to account for the formation of the moon by a different process.

(ii) We have pairs of planets like Venus-Earth, Jupiter-Saturn and Uranus-Neptune. Are they the result of the breakup of larger bodies caused by the tidal action of the sun?

10. Detection of other planetary systems

Direct detection of planets is rather difficult, but it is comparatively easy to detect nebular discs around stars due to their larger size. Both are discussed in a recent article by Beckwith and Sargent (1996).

(a) Direct detection: There are three ways for direct detection of planets: (i) Photometry can be used for observing the eclipses of the star by the accompanying planet. Taking Sun-Jupiter system we need to observe eclipses of $\Delta m = 0^m.01$ with a period of 12 years which is difficult.

However, the planet can also cause increase in brightness by the process of microlensing; this effect is much larger. (ii) Astrometry can show the presence of a planet by the wavy motion of the star. Again for the Sun-Jupiter system we need an accuracy of $\Delta\odot=0''.001$ at 10 AU which might become possible with the Hiparcos satellite. (iii) Spectroscopy can give the variations in the radial velocity of the star caused by its motion around the centre of mass of the star and the planet. Here the accuracy required is $\Delta v_r=12$m/s, it is now possible to measure Δv_r as small as 5 m/s. Four planetary systems are discovered in this way; they are given below:

Star	Mass of the planet	Period	a (AU)
70 Vir	9 M_J	116^d	0.5
47 UMa	3 M_J	1100^d	2
51 Peg	0.5 M_J	......	0.1
55 Cen	0.8 M_J	......	0.2

Variations in the periods of pulsars can also provide evidence for the existence of planets around them. PSR 1257+12 and PSR 1620-26 are found to have planets of mass equal to 2.8 $M_\oplus$ and 3.4 $M_\oplus$, respectively, at 0.5 AU.

(b) Detection of planetary discs: The planetary discs around stars contain dust, so they can be detected by the infrared radiation of the dust. Discs around β Pic and Vega have been detected in this way. Hubble telescope has recently photographed discs around several young stars, which are silhouetted against the bright light of the Orion nebula. And their infrared radiation has beam measured subsequently.

Firstly, we note that it is now realised that disc formation occurs naturally in a collapsing cloud. So young stars show the phenomena of polar winds and equatorial discs. The lifetime of the disc is estimated to be longer than the time required for the formation of the planets. Infrared radiation gives the mass of the dust in the disc. Then assuming a gas/dust ratio of 100 one can get the total mass of the disc. The disc masses are found to lie between 0.0001 to 0.3 $M_\odot$ with majority between 0.01 and 0.1 $M_\odot$ which is estimated to be the mass of the solar nebula.

Some discs have less gas, so they cannot give rise to Jupiter like planets, but several earth like planets will be quite common. Many discs show a central hole like the disc of β Pic, which also shows a tilt indicating the presences of a planet within the system. As the statistics of the planetary discs grow it may be possible to detect the planets in various stages of formation. And then we will have a reliable model for the formation of the solar system.

References

Alfven H., 1954, On the origin of the solar system, Oxford University Press, Oxford.

Beckwith S.V.W., Sargent A.D., 1996, Nature, 383, 139.

Chamberlin T.C., 1916, The origin of the earth, University of Chicago Press, Chicago.

Chandrasekhar S., 1941, ApJ, 93, 285, 323.

Chandrasekhar S., 1946, Rev.Mod.Phys.18,94.

Hoyle F.,1960, Q.J.RA.S.,1,28.
Hughes D.W.,1993, Q.J.R.A.S.,34,401.
Jeans J.H., 1919, Problems of cosmogony and stellar dynamics, Cambridge Univ. Press, Cambridge.
Jeans J.H., 1929, Astronomy and Cosmogony, Dover Pub., New York.
Jeffreys M., 1924, The earth, Cambridge Univ. Press, Cambridge.
Kant E.,1755, Allgemeine Naturgeschifte and theory der Himmels.
Kuiper G.P., 1951, in Astrophysics, Ed. J.S. Hynek, McGraw-Hill, New York, p.357
Laplace P.S., (1796), Exposition du system du monde, Paris.
Moulton F.R., 1905, AJ, 22, 165.
Nölke F., 1930, Der Entwicklungsgang unsers Planetensystems.
Russell H.N., 1935, The solar system and its origin, Mc Millan, New York.
Schatzman E., 1962, Annal d'Astrop, 25,18.
Schmidt O.Y., 1944, Doklady Akademi Nauk, 44,8.
Spitzer L., 1939, ApJ, 90, 675.
Urey H.C., 1952, The planets, Yale Univ. Press.
Von Weizsacker C.F., 1947, Z.f.Phys.,124,614.