Astronomy. — The disc theory of the origin of the solar system. By H. P. BERLAGE, Director Meteorological and Geophysical Service, Batavia.

(Communicated at the meeting of May 29, 1948.)

1. The difficulties raised against the encounter theory of the origin of the solar system, as well as the appearance of two other theories 1) in recent years, may permit the author to review briefly the outcome of a series of papers published in these Proceedings between 1930 and 1940²), including some corrections and extensions to which he found himself induced since. The more so, because he admits that in his odyssae through several attempts to attain a rational picture of the evolution of the planets, he was many times led astray, but is now in the position to formulate a rather concise theory as a working basis.

This theory is essentially monistic, considering the transformation of a nebula rotating round a heavy nucleus. Evidence converges towards this origin since careful analysis proves that also every dualistic conception of the solar system will lead us back to a primeval sun, surrounded by a very extensive gaseous envelope. SPITZER found the right expression, when writing ³):

"Such an atmosphere is reminiscent of the Laplace nebular hypothesis, except that in this case there need be no lack of angular momentum. The validity of the encounter theory as an explanation of the origin of the solar system rests apparently on whether or not a non-uniformly rotating atmosphere could condense into solid bodies."

The author is convinced that this question should be answered in the affirmative.

2. Because the distribution of mass in the solar system is such that only one part in 700 is concentrated in the planets, ROCHE's model of a nucleus surrounded by a massless shell is applicable. Hence, the potential at any point in the atmosphere is Newtonian.

Let *r* denote the distance of a volume element from the axis of revolution, *h* its height above or below the equatorial plane, *p* and ρ gaspressure and

¹) H. ALFVEN, On the cosmogony of the solar system, Stockholms Observatoriums Annaler, Band 14 No. 2, 1942; No. 5, 1943; No. 9, 1946. — C. F. V. WEIZSÄCKER, Ueber die Entstehung des Planetensystems, Zs. f. Ap. 22, 319—355 (1943). Reviewed by G. GAMOV and J. A. HANEK, Ap. J., 101, 249—254 (1945).

²) Proc. Kon. Akad. v. Wetensch., Amsterdam, **33**, 614 (1930); **33**, 719 (1930); **35**, 553 (1932); **37**, 221 (1934); **38**, 857 (1935); **43**, 532 (1940).

³) LYMAN SPITZER Jr., The Dissipation of Planetary Filaments, Ap. J. **90**, 674–688 (1939).

density, M the mass of the sun, γ the constant of gravitation, ω the angular velocity of an element of the nebula, then the two fundamental equations of equilibrium become

$$\gamma M r (r^2 + h^2)^{-\frac{3}{2}} + \frac{1}{\varrho} \frac{\partial p}{\partial r} = \omega^2 r.$$
 (2)

In order to solve these equations we introduce the gas equation

in which T is the absolute temperature, μ the mean molecular weight of the gas and k a universal constant.

It is easily shown that, when the total mass of the planets was originally evenly distributed through a sphere of radius equal to, say, 500 astronomical units, this matter was opaque to solar radiation. Therefore, it is reasonable to assume that temperatures in the solar nebula will never have differed much from black body temperatures. This means a variation of T as the inverse square root of the distance from the sun.

It follows that

when r_0 is the radius of the sun and T_0 the temperature of the nebula where it is in touch with the solar atmosphere.

Let us further assume that the variation of μ can be represented adequately by a certain power of the solar distance.

If

 $s = \frac{1}{2}$ means uniform chemical composition of the nebula.

We then have, when f_0 denotes the value of f at the surface of the sun

Integrating (1) and (2), we obtain

$$\varrho = \varrho_e \left(1 + \frac{h^2}{r^2} \right)^{\frac{1}{2}s} \exp \left[-\frac{\gamma M}{r_0^s f_0(1-s)} \left\{ \frac{1}{r^{1-s}} - \frac{1}{(r^2 + h^2)^{\frac{1}{2}(1-s)}} \right\} \right] \quad . \tag{7}$$

$$\frac{\gamma M}{r^2} + f_0 \left(\frac{r_0}{r}\right)^s \frac{d \log \varrho_e}{d r} - s f_0 \frac{r_0^s}{r^{s+1}} = \left(1 + \frac{h^2}{r^2}\right)^{\frac{1}{s}} \omega^2 r \quad . \quad . \quad (8)$$

in which ρ_e denotes the density of the gas in the equatorial plane. If it is given the two formulae define the structure of the nebula in steady motion.

Let us consider briefly 3 models which have special features, viz.:

b)
$$s = \frac{1}{2}$$
, $\varrho = \varrho_e \left(1 + \frac{h^2}{r^2} \right)^{\frac{1}{2}} \exp \left[-\frac{2\gamma M}{r_0^{\frac{1}{2}} f_0} \left\{ \frac{1}{r^{\frac{1}{2}}} - \frac{1}{(r^2 + h^2)^{\frac{1}{2}}} \right\} \right]$. (7b)

$$\frac{\gamma M}{r^2} + \int_{0}^{0} \left(\frac{r_0}{r}\right)^{\frac{1}{2}} \frac{d \log \varrho_e}{dr} - \frac{1}{2} \int_{0}^{0} \frac{r_0^{\frac{1}{2}}}{r^{\frac{3}{2}}} = \left(1 + \frac{h^2}{r^2}\right)^{\frac{1}{2}} \omega^2 r \quad . \quad (8b)$$

c)
$$s=1$$
, $\varrho=\varrho_e\left(1+\frac{h^2}{r^2}\right)^{\frac{1}{2}}\exp\left[-\frac{\gamma M}{2r_0f_0}\left(\frac{h}{r}\right)^2\right]$ (7c)

$$\frac{\gamma M}{r^2} + f_0 \frac{r_0}{r} \frac{d \log \varrho_e}{dr} - f_0 \frac{r_0}{r^2} = \left(1 + \frac{h^2}{r^2}\right)^{\frac{1}{2}} \omega^2 r \quad . \quad . \quad . \quad (8c)$$

a) s = 0 means a decrease of the mean molecular weight of the gas with solar distance proportional to the decrease of temperature. The resulting structure is the simplest of the three, e.g. ρ decreases monotonously with increasing h. The model is the only one showing the property that ω is a function of the distance r from the axis of rotation only.

b) $s = \frac{1}{2}$ means uniform chemical composition of the nebula. Many will claim the greatest a priori probability for this case. The equilibrium structure shows, however, one unreasonable feature. Lim $\rho = \infty$ for $\lim h = \infty$.

c) s = 1 means an increase of the mean molecular weight of the gas with distance from the sun inversely proportional to the decrease of temperature. The infinite increase of μ with solar distance cannot be strictly true, of course, but the limiting value of ϱ is now zero for lim $h = \infty$. Another satisfactory feature of this structure is that lim $\omega = 0$ for lim $h = \infty$. This too is valid for model b), but model c) is simpler. If ϱ_e is the same all over the equatorial plane, a meridional section through the nebula shows straight lines of equal density intersecting at the centre.

Now, the premises are such that clearly these 3 models encompass all reasonable possibilities. The real structure will have shown characteristics somewhere intermediate between the characteristics of the structures just described and probably not have differed much from model b). When the author proposes to proceed this investigation with model a), it is principally because the calculations in the next paragraphs are feasible only when the nebula rotates cylinderwise. Consequently all further steps taken are on surer ground. Yet, there might be more plausibility in assumption a) than seems to be the case.

For even if the chemical composition of the nebula was originally

homogeneous, it is very improbable that this was maintained. The molecules did not move in free orbits. Gravitation was partly supported by gaspressure. Hence, the heavier molecules will have shown a tendency to gather towards the centre, the lighter ones towards the periphery.

Many will point out that the very high percentage of H and He present in the original nebula kept the mean molecular weight of the gas at some low value, say 4, all over. The actual constitution of the planets, however, is incompatible with such a large dosis of H and He and when the planets were not able to retain the H and He, it is extremely probable that even if H and He were once present in the nebula in the amount in which they are present in the sun now, they vanished from the nebula by evaporation into space before the planets were formed.

As to the probable existence of molecules in the nebula, we may mention that comets tails of the same low density consist of several molecular compounds of H, C, N, and O of which many show signs of disintegration only at small solar distances. We might even try to explain the general trend of the actual densities of the planets by a preponderance of molecules of heavier or less heavy weight where the planets condensed. There is at least no danger of losing any essential aspect of the development of the planets, when we go as far as the very simple assumption $f = f_0 =$ constant. Quantitatively model a) works out as follows.

We know that Fe = 56 and $SiO_2 = 60$ were predominant building materials of the nucleus and the outer shells of the earth. This would mean $\mu = 58$ at the solar distance where the earth was formed. The blackbody temperature at this same distance being 290°, we obtain $f_0 = 4 \times 10^8$ c.g.s. to base our further calculations upon.

3. For $h = \infty$ the density of the gas tends towards the finite value

$$\lim_{h=\infty} \varrho = \varrho_e \exp\left[-\frac{\gamma M}{r}\right] \dots \dots \dots \dots \dots (9)$$

The solar nebula behaved morphologically like a vortex in the interstellar medium, as DESCARTES conceived it. As the interstellar medium established the external boundary conditions, the fact that (9) depends on r means, of course, that the motion of the nebula could not be strictly stationary. Independent internal boundary conditions are established by the sun at its surface. We have to consider the nebula as being in slow but gradual development, its structure being an equilibrium structure only in first approximation.

For $h \ll r$, (7a) reduces to

$$\varrho = \varrho_e \exp\left[-\frac{\gamma M h^2}{2f_0 r^3}\right] \cdot \cdot \cdot \cdot \cdot \cdot \cdot \cdot (10)$$

Hence, with $\gamma = 6.67 \times 10^{-8}$ and $M = 2.00 \times 10^{33}$, we obtain roughly

when r and h are expressed in astronomical units.

In order to fix the ideas let us consider two instances. The density of the gas drops to a value 10^{-10} times the density in the equatorial plane at a vertical distance as small as h = 0.1 r for r = 5 astronomical units, that is as far out as Jupiter, whereas the drop is to one hundredth under the same conditions at r = 25 astronomical units, that is halfway Uranus and Neptune.

In other words, if these conditions of temperature and composition have prevailed in the solar nebula it has been a thin flat disc over most of its extent. The models b) and c) are for given total mass of the nebula less flat in the inner portions, but flatter in the outer portions, when compared with model a).

The theory of a gaseous envelope of the sun in steady motion and its flat disc-like shape was developed later independently by v. WEIZSÄCKER.⁴ He uses it at the base of an extremely interesting hypothesis of the origin of the planetary system, which is, however, fundamentally different from the hypothesis developed here.

4. The final structure of the disc is determined by viscosity. It must have acted first rapidly but roughly as turbulent viscosity, later, where turbulence dies out, slowly as the ordinary viscosity of laminar motion.

It was shown in Proc. 38 that the theorem of minimum loss of energy by viscosity requires, when squares and products of small quantities are neglected, that

The solution of this equation is

$$\varrho_e = \varrho_0 \exp[-ar^{\frac{1}{2}}]$$
 (13)

when ρ_0 and *a* are integration constants to be determined from observational data. These are the constant mass and the constant moment of momentum of the disc, or

$$m = 2.26 \times 10^{30} \text{ gr}$$

 $\theta = 3.14 \times 10^{50} \text{ c.g.s.}$

In the way indicated in Proc. 35, we obtain

$$a = 6.82 \times 10^{-7} \text{ cm}^{-\frac{1}{2}}$$

 $\rho_0 = 4.63 \times 10^{-9} \text{ gr cm}^{-3}$

We have now reached the point, where we can draw a picture of the nebula. Its meridional section is shown in Fig. 1. Lines of equal density suggest a toruslike structure. A 10^{-24} surface would indicate where the nebula merges into interstellar matter, since the latter is found to possess this order of density within the galactic system. A 10^{-22} line has been dotted, because even this hardly stands for the boundary surface. The 10^{-20} surface should be regarded as the most rational envelope of the original nebula. A density of 10^{-18} is reached in the earth's atmosphere

at a height of 300 km, where we know the atmosphere is dense enough to cause it start visible phenomena like polar lights and the radiation of

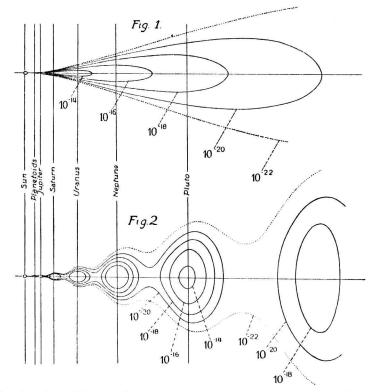


Fig. 1 and 2. Meridional section through gaseous disc generating planets.

meteorites. The surface 10^{-18} still envelopes Pluto's orbit. It might be argued that the matter between the envelopes 10^{-18} and 10^{-20} became the stock of our comets.

5. The motion described so far is steady. According to Proc. 35 it is also stable, where the following condition is satisfied,

$$\gamma M + f_0 \frac{d}{dr} \left(r^3 \frac{d \log \varrho_e}{dr} \right) > 0. \quad . \quad . \quad . \quad . \quad (14)$$

Applying this condition, we find that there is a critical radius

such that the nebula will have been stable and in laminar motion up to a distance from the axis of rotation which is smaller than r_c , whereas the nebula will have been instable and in turbulent motion in its outer parts. Inserting the values of a and f_0 found before, we obtain

$$r_c = 5.31 \times 10^{15} \text{ cm}.$$

As this distance is far beyond Pluto's orbit, the solar nebula may be supposed to have settled down in laminar motion up to its outskirt.

As was shown in Proc. 35 *a* is independent of f_0 . Hence, even if the molecular weight of the gas had been only one tenth of the weight which was assumed before, we find

$r_c = 1.14 \times 10^{15} \, \mathrm{cm}$

which proves that the nebula would still have been stable up to a distance beyond Pluto's orbit.

There seems to be no theoretical reason for large scale convective motions of the disc such as are at the basis of WEIZSÄCKER's theory of the origin of the planets and certainly even less so when WEIZSÄCKER's assumption that the particles move in free orbits would prove to be of vital significance. However, observational evidence favours the hypothesis that the motion of the disc actually was or at least remained turbulent down to a shorter distance from the sun than the theoretical limit. As a matter of fact, the motion of Pluto shows already signs of being less strictly regulated than the motion of the other planets, while the turbulent motion of the girdle of the nebula beyond Pluto might explain the retrograde motion of many comets.

6. So far we did develop the theory of the rotating solar nebula up to first order small quantities and found it steady within these limits. We know, however, that it cannot attain a strictly steady state and proceeds in continuous evolution. Hence, we are confronted with the problem whether the nebula will condense into planets or keep the disclike structure described so far. From Proc. **37** we know that steady state and KEPLER motion are only compatible, if the coefficient of viscosity would have adjusted itself throughout by adequate migration of molecules. The more probable method of compromise, however, is a dividing up of the nebula in concentric zones in each of which KEPLER's third law holds approximately whereas the radial gradient of the angular velocity shows discontinuities from zone to zone.

It is easily shown that this adjustment would result in the generation of circular zones in which the angular velocity is successively superior and inferior to the KEPLER velocity, which in the steady state means a fluctuation of the density ϱ_e with solar distance, that is a tendency towards concentric ring formation.

We may add that the formation of separate rings is the natural way in which the loss of energy by friction can decrease further and would eventually vanish when the rings rotate like rigid bodies. Thus fortified in our assumption that the planetary system once passed through the state of concentric gaseous rings, let us consider a configuration of the nebula which is infinitely near to the one described by (13), but differing from it by the addition in the exponent of an ondulating term. We than get

$$\varrho_e = \varrho_0 \exp\left[-a r^{\frac{1}{2}} + \varepsilon \sin f(r)\right] \quad . \quad . \quad . \quad . \quad . \quad (16)$$

the amplitude of the ondulation being an infinitesimal dimensionless quantity. As (13) was the solution of (12), the only condition which has to be satisfied by f(r) is that the left side of (12) when multiplied by r, remains smaller than \varkappa for every value of r, when \varkappa denotes an infinitesimal dimensionless quantity.

Now, it is easily shown that among the functions expressing some power of r, the logarithmic function

$$f(\mathbf{r}) = p \log \mathbf{r} + q \quad \dots \quad \dots \quad \dots \quad \dots \quad (17)$$

replacing the zero power is the only function which keeps the left side of (12) small for any value of r. This may be considered as an indication that whenever there is a tendency of the nebula to change from a disc into a series of concentric rings the radii r_n and r_{n-1} of two successive rings will satisfy the relation

This, however, is the expression of BODE's rule, when we disregard the additional term 4).

The mass of a planet m_p can be expressed by the formula

$$m_p = \text{constant} \times r^{\frac{5}{2}} \exp\left[-a r^{\frac{1}{2}}\right] \bigtriangleup r \quad . \quad . \quad . \quad . \quad (19)$$

when $\triangle r$ denotes the difference between two successive radii, where minimum density occurs. As, however, following (18)

when r_p is the planets distance from the sun, we find

$$m_p = \operatorname{constant} \times r_{p^2} \exp\left[-a r_{p^2}\right]$$
. (21)

This explains the general trend of the masses of the planets, which is increasing first and decreasing again, when we pass through the system from the centre towards the periphery.

Strictly, of course, when developing any arbitrary oscillation, we have to add in the exponent of (16) several terms

$$\varepsilon_1 \sin (p_1 \log r + q_1) + \varepsilon_2 \sin (p_2 \log r + q_2) + \dots \quad (22)$$

Observational data, however, favour the assumption that one of the partial terms is naturally stressed in the act of concentration and accounts for BODE's rule. The other terms account for the deviations of the distances of the planets from BODE's rule, as well as for the deviations of the masses of the planets from the general trend just mentioned.

The problem can also be treated as a problem of gravitational instability. In our nebula showing axial symmetry the tidal forces of the central body, which in other circumstances would tend to disintegrate local condens-

$$\sqrt{r_n} - \sqrt{r_{n-1}} = \text{constant}.$$

⁴⁾ The reader will have noticed that by erroneous reasoning the author was led in Proc. 43 to an alternative kind of density variation and to the following rule of distances

ations, are balanced by rotation. Therefore, its condensation into rings cannot differ much from what would happen, when a non-rotating disc of the same structure, but without central body, would be left to itself at a given moment. It is essentially the same process. The author even tried whether the variation of the intensity of the solar radiation with the sunspot cycle might have been the source of a pressure wave spreading through the disc lifting it over the threshold of gravitational instability and found that it would lead to the right dimensions.

7. Let us now prove that a variation of the course of the density from (13) to (16) may lead to a natural increase of the amplitude of the fluctuations introduced. In order to do this, we have to calculate the variations of mass δm , of angular momentum $\delta \theta$, of potential energy δV and of kinetic energy δU , involved in the density variation. Let the conditions

$$\delta m = 0$$
, $\delta \theta = 0$, $\delta (U+V) = 0$... (23)

be satisfied. When these conditions are compatible with the condition

$$\delta U = -\delta V > 0$$
 (24)

we may assume that the disc follows this course and tends to change into a series of concentric rings.

In Proc. 43 the author gave affirmative evidence. There he assumed, however, that the nebula before becoming unstable, passes through a state of equilibrium in which $\delta U = 0$ and $\delta V = 0$ are satisfied up to first order small quantities automatically, confining himself to the calculation of second order small quantities. Apparently, this assumption is not allowed. Using the formulae developed in Proc. 43 the author was able to prove that condition (24) reads

$$\int_{0}^{\infty} r^{\frac{1}{2}} \exp\left[-a r^{\frac{1}{2}}\right] \sin\left(p \log r + q\right) dr < 0. \quad . \quad . \quad (25)$$

It does not invalidate the conclusion reached in Proc. 43. As (25) is independent from the other conditions and the left side an oscillating function, it can be satisfied.

The natural tendency leading in the direction of the formation of rings, we are tempted to follow the process of condensation through finite values of ε to its critical value. From the actual dimensions of the planetary system, we derive

$$p = 10.86$$

when 10 is the base of the logarithm in (25). Final condensation will occur when the density in the densiest part of every ring outgrows the minimum value stated in ROCHE's theorem, that is when

$$\varrho_0 \exp\left(-a r^{\frac{1}{2}} + \varepsilon\right] > 14.5 M \left(\frac{4}{3} \pi r^3\right)^{-1}$$
. (26)

When ε_c is the oritical amplitude of the density wave and known values of *a* and ϱ_0 are inserted, we get

$$\varepsilon_c = (42.17 - 3^{10}\log r) ({}^{10}\log e)^{-1} + 6.82 \times 10^{-7} r^{\frac{1}{2}}$$
. (27)

805

From this relation we obtain the following table

r = cm	ε
1012	14.9
1013	7.3
1014	7.2
1015	15.0

 ε_c has a very flat minimum at $r = 7.74 \times 10^{13}$ cm, which is very near Jupiters mean distance from the sun. The lowest value of ε_c is 7.15. This proves that most probably condensation proceeded from Jupiter towards Mercury and from Jupiter, towards Neptune, where ε_c reaches the values 10.6 and 10.3 respectively. The critical density ϱ_c in the equatorial plane which has to be surpassed is greatest for Mercury, where it reaches the value

$$\varrho_c \equiv 3.50 imes 10^{-5} \mathrm{~gr~cm^{-3}}$$

In the nearest density throughs the density is then of the order 10^{-14} gr cm⁻³. This shows how far the development towards individual rings has gone. It is reasonable to assume that the final condensation started with $\varepsilon = 8$. Fig. 2 shows how the meridional section through the disc looks like when the evolution has progressed up to this point.

We are still confronted with two questions. The first is whether in this stage kinetic energy can still grow at the cost of potential energy. We are allowed to answer it in the affirmative, since we will be able to show that there are cases in which the final system of condensed globes possesses more kinetic energy than the original nebula. The second question is, whether the motion of the gas could remain laminar up to the point of final condensation.

Introducing (16) in (14), we get the following condition of stability,

$$\frac{\gamma M}{f_0} - \frac{5}{4} ar^{\frac{5}{4}} + \varepsilon pr \{2\cos(p\log r + q) - p\sin(p\log r + q)\} > 0. \quad (28)$$

The maximum value attained by the function between brackets is 11.25. When the other known values are inserted, we obtain

From this relation the following table can be derived

r =	cm	ε
	1012	2720
	1013	272
	1014	27.2
	1015	2.5
5.31 $ imes$	1015	0

As was already shown in 5., ε_c decreases to zero at a distance of

 5.31×10^{15} cm from the sun. Beyond this distance the solar nebula would be in turbulent motion previous to any tendency of condensation. Both critical limits are equal at 3.02×10^{14} cm that is slightly beyond Uranus, where they reach the value 8.9. This proves that all the rings may have remained stable with the exception of the rings of Neptune and Pluto. They reached the limit of turbulent motion before reaching the limit of final condensation. According to our assumption this means that Neptune and Pluto were in danger to remain in the form of planetesimals. Neptune has evidently escaped this danger. Pluto perhaps not. Its excentric behaviour suggests that there may be other "plutoids".

Jupiter, when first starting to condense, may have drained his surroundings from matter before the other planets came into existence. This may be the reason, why there was almost no matter left for a planet in the place where we actually find the planetoids. The planetoids never gathered into one planet probably because matter in the ring was too dispersed.

Saturn's ring remained in the planetesimal state as it revolves inside ROCHE's limit. On the other hand the low density of Saturn's satellite Mimas is only compatible with a planetesimal constitution like a comet's head, which proves that the final concentration of one globe from planetesimals is a possibility. Perhaps even most bodies in the solar system passed through the state of a cluster of meteorites. These considerations show that CHAMBERLIN and MOULTON's planetesimal theory may be essentially right, as the theories of DESCARTES, KANT and LAPLACE proved also essentially right.

What happens to the planetesimals, building up comet's heads, when they come too near to the sun and evaporate might be the opposite from the condensational process which once started in colder portions of the nebula.