FORMATION OF THE GIANT PLANETS†

D. J. STEVENSON

Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, CA 91125, U.S.A.

Abstract—Observational constraints on interior models of the giant planets indicate that these planets were all much hotter when they formed and they all have rock and/or ice cores of ten to thirty earth masses. These cores are probably soluble in the envelopes above, especially in Jupiter and Saturn, and are therefore likely to be primordial. They persist despite the continual upward mixing by thermally driven convection throughout the age of the solar system, because of the inefficiency of double-diffusive convection. Thus, these planets most probably formed by the hydrodynamic collapse of a gaseous envelope onto a core rather than by direct instability of the gaseous solar nebula. Recent calculations by Mizuno (1980, *Prog. Theor. Phys.* **64**, 544) show that this formation mechanism may explain the similarity of giant planet core masses. Problems remain however, and no current model is entirely satisfactory in explaining the properties of the giant planets and simultaneously satisfying the terrestrial planet constraints. Satellite systematics and protoplanetary disk nebulae are also discussed and related to formation conditions.

INTRODUCTION

Most discussions of giant planet formation are prejudiced by some specific set of assumptions concerning the relevant processes and sequence of events during formation. I hope to avoid this by emphasizing the observational constraints: what can be said about the present state and evolution and what are the implications for formation? In fact, there is no formation scenario yet constructed which satisfies all of the constraints discussed here. There exist models which might satisfactorily account for the observed properties of the planets but which raise a number of unresolved dynamical issues concerning the early state of the solar system. Conversely, there exist models which are dynamically plausible but which are not obviously compatible with the inferred internal properties of the giant planets. The synthesis is not imminent but the problem is well posed.

I begin with a discussion of interior models of the giant planets, emphasizing that all these planets (Jupiter, Saturn, Uranus and Neptune) have essentially adiabatic hydrogen–helium envelopes and central regions ("cores") of rock

and/or ice of ten to thirty earth masses. The remarkable similarity of the core masses may be profoundly significant and occurs despite the wide range of total planetary mass (from 318 $M_\oplus$ for Jupiter to 14.6 $M_\oplus$ for Uranus; 1 $M_\oplus$ = one earth mass). The adiabatic envelopes imply high internal temperatures and impose constraints on the thermal evolution. Initially hot states are implied for all the giant planets, although the implication is most direct and compelling for Jupiter and Saturn. The possibility of compositional evolution is discussed and it is shown that the present cores are most probably primordial. Although their constituents are likely to be soluble in metallic hydrogen, mixing upwards by thermal convection is shown to be inefficient.

The satellite systems are discussed as potential "tracers" of the thermal conditions in the circumplanetary disk of gas and dust during accretion. This is probably important for Jupiter but difficult to apply to Saturn and not yet useful for the other satellite systems.

With these observational constraints in mind, two sections are devoted to formation scenarios: one on the purely gaseous instability hypothesis (such as Jeans collapse) and one on the nucleated instability hypothesis, in which a core of rock and/or ice initiates hydrodynamic collapse of the surrounding gaseous envelope. The former is dynamically pleasing because it ensures very rapid formation (desirable, at least for Jupiter) but it does

not immediately explain the similarity of core masses for the giant planets. The nucleated instability is delayed by the accretion time of the rock/ice core but may explain the core masses. I conclude with some suggestions for the required synthesis.

INTERIOR MODELS

Figure 1 and Table 1 together summarize the most important constraints and inferences for giant planet interiors. The figure shows the mass–radius relationship for bodies of various compositions, calculated from existing shock wave data and theoretical equations of state for cosmically abundant materials (Stevenson, 1982). The solid lines are for $T = 0 \, \text{K}$ bodies whereas the dashed lines are for adiabatic bodies with specific entropies chosen to be appropriate for the present thermal states of these planets. The superimposed positions of the giant planets clearly indicate that Jupiter and Saturn are primarily hydrogen–helium but with Saturn deviating more from a cosmic mixture than Jupiter. In fact, the more nearly horizontal dashed line for the hydrogen–helium adiabatic body, which arises because $P \propto \rho^2$ is a good first approximation to the equation of state (P is the pressure and ρ is density), means that Saturn would have almost the same radius as
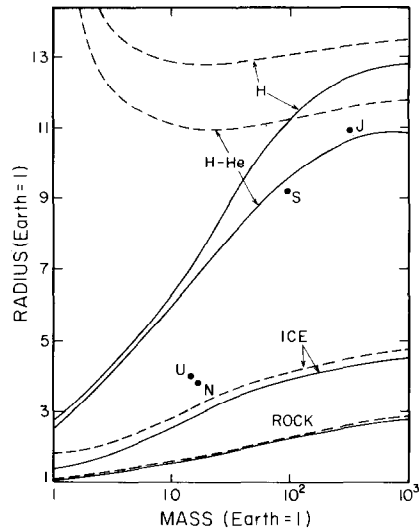

FIG. 1. THE MASS–RADIUS RELATIONSHIP FOR SELF-GRAVITATING BODIES. THE SOLID LINES ARE FOR COLD MATTER ($T = 0 \, \text{K}$) AND THE DASHED LINES ARE FOR ISENTROPES CHOSEN TO BE APPROPRIATE FOR THE PRESENT JUPITER OR SATURN IN THE CASE OF H AND H–HE, AND APPROPRIATE FOR URANUS AND NEPTUNE IN THE CASE OF "ICE" (A COSMIC MIXTURE OF H_2O, NH_3 AND CH_4) AND "ROCK". The curves labeled H–He are for a 75% H–25% He mixture by mass. Notice the insensitivity of radius to mass for isentropic bodies of H or H–He. The positions of the four giant planets are indicated by J, S, U and N.

TABLE 1*

	Jupiter	Saturn	Uranus	Neptune
Mass (Earth = 1)	318.05	95.147	14.58(±0.1)	17.23(±0.08)
Volume-averaged radius (10^9 cm) (1 bar level)	6.980 ± 0.001	5.830 ± 0.003	2.55 ± 0.02	2.45 ± 0.03
Average density (g cm^{-3})	1.334 ± 0.006	0.69 ± 0.01	1.26 ± 0.07	1.67 ± 0.1
I/MR^2	0.250 ± 0.005	0.23 ± 0.01	0.16 ± 0.02	0.13 ± 0.03?
Rock + ice core mass (earth = 1)	10–30	15–25	11–31	13–15
Excess luminosity (erg s^{-1})	$(4 \pm 1) \times 10^{24}$	$(1.3 \pm 0.3) \times 10^{24}$	$\lesssim 1.5 \times 10^{22}$	$(2 \pm 1) \times 10^{22}$
Central temperature (approx)	$2 \times 10^4 \, \text{K}$	$1.2 \times 10^4 \, \text{K}$	$\sim 6 \times 10^3 \, \text{K}$	$\sim 7 \times 10^3 \, \text{K}$

*See Stevenson (1982) for more details.

Jupiter if it had the same composition. The small radius of Saturn is explained by a greater enrichment of heavier elements in the form of "rock" (silicates, iron) and "ice" (H_2O, NH_3, CH_4) relative to cosmic abundance. Figure 1 does not provide an immediate interpretation of Uranus and Neptune but suggests that these bodies are ice-rich.

The gravity field supplies additional information on the composition and possible layering of these oblate planets. The moment of inertia coefficient I/MR^2 is 0.4 for a uniform density sphere, 0.331 for the earth, 0.26 for a homogeneous sphere in which $P \propto \rho^2$, 0.25 for Jupiter and decreasing monotonically as one moves out in the solar system. These small values for the giant planets are all indicative of substantial cores or central concentrations of material that is much more dense than hydrogen–helium at the same pressure. Table 1 lists the inferred core masses of the giant planets, all of which are within a factor of two of 15 $M_\oplus$. The modeling procedures responsible for these inferences are described fully elsewhere (Stevenson, 1982).

The thermal balance of these planets is a very important constraint on the evolution. In each case except Uranus, the emitted i.r. is about twice the absorbed insolation. The similarity of the factor of two is only coincidence; it is more important to compare this energy output with the available thermal energy of the planet. One hypothesis (equivalent to Lord Kelvin's erroneous attempts to explain the heat outflows of the Sun and the Earth) is that gradual, secular cooling from an initially much hotter state may explain the present internal heat sources of these planets. To test this hypothesis, it is first necessary to determine the thermal structure. Fortunately, this is greatly simplified by the inequality

$$-k\frac{dT}{dr}\bigg|_{ad} \ll F, \qquad (1)$$

where k is the thermal conductivity (including radiative transport), the adiabatic temperature gradient is given by $(dT/dr)|_{ad} \equiv -\alpha Tg/C_p$ (α is the coefficient of thermal expansion, g is the gravitational acceleration and C_p is the specific heat) and F is the actual heat flow. This inequality is usually satisfied by about at least a factor of ten at present and even more strongly in the past. It implies that microscopic processes (radiation, conduction) are incapable of transporting the observed heat flow from deep within the planet along a convectively stable (i.e. subadiabatic) temperature gradient.

Convection ensues and the resulting temperature structure is almost exactly adiabatic. (Convection requires only a small superadiabaticity because the adiabat does not encounter any solid–fluid phase boundaries, except perhaps in the core. All these planets have bottomless atmospheres.)

The adiabatic structure greatly simplifies consideration of the thermal evolution. To a good approximation, the luminosity L of a homogeneously evolving, degenerate body (one in which there is no redistribution of constituents and little change in radius during the evolution) is given by

$$L \equiv 4\pi R^2 \sigma(T_e^4 - T_0^4) \simeq -\frac{d}{dt}[M\bar{C}_v T_i], \qquad (2)$$

where σ is a Stefan–Boltzmann's constant, T_e is the effective temperature, T_0 is the equilibrium effective temperature in the absence of an internal heat source (non-zero because of the presence of the Sun), $\bar{C}_v$ is the average specific heat per gram, T_i is an appropriately defined internal temperature and M, R are the planetary mass and radius. For an adiabatic planet, $T_i \propto T_e$ and it follows that the time τ elapsed for the planet to cool from an initial effective temperature T_{ei} to the present value T_{eo} is

$$\tau = \frac{M\bar{C}_v \vartheta}{4\pi R^2 \sigma} \int_{T_{eo}}^{T_{ei}} \frac{dT}{T^4 - T_0^4}, \qquad (3)$$

where $\vartheta \equiv T_i/T_e$ is a constant. The precise choice of the upper limit of the integration T_{ei} is unimportant if $T_{ei} \gtrsim 2T_{eo}$ because of the T^{-4} dependence of the integrand. Evaluation of τ is discussed by Hubbard (1980). For Jupiter, it is indistinguishable from the age of the solar system. This suggests a very simple and physically plausible explanation for Jupiter's excess luminosity: secular cooling from an initial hot state. Saturn's excess luminosity is too large to be explained in this way (the calculated value of τ is $\sim 2.5 \times 10^9$ years, smaller than the age of the solar system). If we exclude the implausible hypothesis that Saturn formed much more recently than Jupiter then some additional energy source is required. The most likely explanation is the formation of helium raindrops deep within the planet because of the limited solubility of helium in metallic hydrogen (Stevenson, 1980). Downward migration of these raindrops releases gravitational energy as heat. It should be noted that this process does not explain all the heat output—it is still necessary to have a "hot start" ($T_{ei} \gg T_{eo}$). Uranus has no detectable

excess luminosity (but this does not mean that it has none, see the discussion in Hubbard, 1980). Interestingly, Neptune's excess luminosity is lower than would be expected for homogeneous cooling from a "hot start" ($T_{ei} \gg T_{eo}$). This could conceivably be explained by a "cold start" ($T_{ei} \simeq 1.5T_{eo}$) but the gravitational energy of formation is so large that this is very unlikely. A more plausible explanation is that gravitational work has been done redistributing stably layered but soluble constituents.

Redistribution of constituents can occur in two ways. First, convection can redistribute soluble constituents from an initially stably stratified situation. Thus, if the planet formed by collapse of a gaseous envelope onto a rock/ice core, then convection in the envelope could subsequently "erode" the core. This always causes a reduction of heat output. Second, constituents can be redistributed if they have limited solubility. Thus, a planet could form a core by "rain out" starting from an initially homogeneous state. This would always cause energy release. Analysis of likely phase diagrams (Stevenson and Salpeter, 1977a; Stevenson, unpublished) indicate that only helium and water are likely to be significantly insoluble. Furthermore, helium insolubility is only important in Saturn and the limited solubility of water in molecular hydrogen is only important in Uranus and Neptune (Stevenson and Fishbein, 1981). The "erosion" of a rock core by convection is therefore a likely process, at least in Jupiter and Saturn.

A lengthy digression into the condensed matter physics which determines solubilities is not appropriate for this paper and will be presented elsewhere, but a few general comments may aid the reader to understand this important issue. At very high pressures ($P \gtrsim 10$ Mbar), the distinction between ionic, covalent and metallic bonding is blurred and the major determinant of solubility is the difference in cell boundary electron density between pure solute and pure host. If this difference is large then the microscopic "surface energy" between host and solute is large and the corresponding solubility is low. Helium not only has one of the largest electron density differences relative to hydrogen in the pressure range of interest, but it is also the most abundant minor species. Consequently, it is the most likely to undergo unmixing in a metallic hydrogen solution. In fact, this unmixing does not occur in Jupiter and is only marginally possible in Saturn (despite its importance to the energy budget, see Stevenson, 1980).

Calculation of the redistribution of stably stratified constituents in a convective planet is a difficult dynamic problem. Two limiting cases can be conceived. A rigorous upper limit to the degree of redistribution can be obtained by appealing to the second law of thermodynamics. The application of this law is not trivial in a convective system (Hewitt et al., 1975) but for our purpose, the inequality can be written approximately as

$$W_{\text{Grav}} < F_B, \qquad (4)$$

where W_{Grav} is the work done redistributing heavy constituents and F_B is the buoyancy flux, given by $\rho g \alpha \langle \Delta T v \rangle l$, where ρ is the fluid density, g is the gravitational acceleration, α is the coefficient of thermal expansion, l is the vertical lengthscale of the convective motions and $\langle \Delta T v \rangle$ is the average of the product of temperature fluctuations and vertical convective velocity. The heat flux $F \equiv \rho C_p \langle v \Delta T \rangle$, where C_P is the specific heat. If we define the temperature scaleheight $H_T \equiv C_p / \alpha g$ then $F_B \simeq (l/H_T)F$. In the situations of interest, $l \sim H_T$ and $F_B \sim F$. (In the conventional language of heat engines, we are dealing with systems in which the temperature drop is very large and the Carnot efficiency is correspondingly high.)

If we are redistributing a constituent which is much more dense than hydrogen then $W_G \simeq (1/2)M_x g l$ where M_x is the mass flux and the material is uniformly redistributed through a vertical height l from an initially localized source. The total redistributed mass is then $M_x < \int (2F/gl) \, dt$. Allowing for the time-dependence of F given by equation (2), one finds that the upper bound to M_x is comparable to the mass of the planet! (Most of the redistribution, if it took place, would have occurred in the early evolution when the heat flow was much larger.) This result is not surprising because of the high Carnot efficiency and the virial theorem (which ensures that the total integrated heat outflow after formation is a substantial fraction of the gravitational binding energy of the planet).

More realistically, the redistribution is partially inhibited by diffusive processes. Studies of convective systems in which the buoyancy effects of heat and composition are opposed indicate that uniformly mixed layers form separated by thin interfaces in which diffusion limits transport (Turner, 1973). In these circumstances, $W_G \sim (D/K)^{1/2}F_B$ where D is the solute diffusivity and K is the thermal diffusivity (Linden and Shirtcliffe, 1978). For $D \sim 10^{-3} - 10^{-4} \text{ cm}^2 \text{ s}^{-1}$ $K \sim$

10^{-1} cm^2 s^{-1} (metallic hydrogen) perhaps a few percent of the planetary mass can be redistributed from a stable post-formation configuration. For $K \sim 10^{-3} - 10^{-2}$ cm^2 s^{-1} applicable to molecular hydrogen, the fraction can be larger. (This is more relevant for Uranus and Neptune where there is no metallic hydrogen.) The calculation is only suggestive since the $(D/K)^{1/2}$ factor is strictly only applicable to steady state and a larger W_G could be sustained in an evolving system. Nevertheless, it suggests an explanation for the partial preservation of the dense cores in these planets inferred from the moments of inertia. Evidently, these cores are most likely primordial. Indeed, they are probably responsible for the formation of the giant planets. This hypothesis is elaborated below in the discussion of nucleated instabilities.

SATELLITES AND PROTOPLANETARY NEBULAE

Each of the giant planets except Neptune possesses an extended family of regular satellites. These satellite systems provide important constraints on planetary formation and early evolution. (It is not clear whether Neptune is indeed an exception, but the unusual orbit of the large satellite Triton indicates that caution must be used in inferring information about Neptune's origin and evolution from its satellites.) Aside from the regularity of orbits, the most striking feature is the density gradient within the Galilean satellites from "rock" ($\bar{\rho} \simeq 3.5$ g cm^{-3}) for Io to the ice-rich Callisto ($\bar{\rho} \simeq 1.8$ g cm^{-3}). The implication is that these satellite systems were most probably formed as "miniature solar systems" by condensation and accumulation of solid matter within protoplanetary nebulae.

There are several intriguing similarities of these hypothetical protoplanetary nebulae. Consider, for example, the minimal reconstituted nebula for each of these planets, defined as the total mass of solar composition material sufficient to provide the observed satellite masses and compositions by condensation. In each case, the minimal reconstituted nebula is within a factor of two of being 4% of the mass of the planet. (Curiously, an analogous result is obtained for the reconstituted solar nebula.) It is also found that the total angular momentum in each reconstituted nebula is equal (to within a factor of two) to the angular momentum of the planet. (In this case, the analogous solar nebula result is totally different, even if one allows for a rapidly rotating early sun.) These similarities are not yet explained, but suggest a similar for-

mation process for each giant planet. I shall focus here on the Jovian system, where the constraints on the nebula are strongest, but will also discuss briefly the Saturnian system.

If a nebula of 0.05–0.1 M_J were uniformly spread over a circular area of radius 30 R_J (M_J, R_J being the present mass and radius of Jupiter) then an average surface density $\bar{\sigma} \sim 10^6$ g cm^{-2} results. At a temperature of ~ 200 K (plausible for the formation region of Ganymede), the scaleheight and mass density of the nebula are $\sim 10^{10}$ cm and $10^{-4} - 10^{-5}$ g cm^{-3}, respectively. Notice that since the scaleheight is small compared to the radial extent, the nebula is actually a disk. The opacity of molecular hydrogen at this temperature and density is $\sim 10^{-4} - 10^{-5}$ cm^2 g^{-1}. It follows that the nebula is optically thick in molecular hydrogen alone, both radially and in the direction vertical to the nebula midplane. It is reasonable to suppose that such a nebula is adiabatic or near-adiabatic.

If the nebula has the same adiabat as the interior of the newly formed or forming Jupiter, and if this adiabat corresponds to temperatures twice as high as the present internal temperature (needed to explain the present heat output of Jupiter), then

$$T \simeq 17{,}000 \, \rho^{0.44} \, K. \tag{5}$$

This is discussed in more detail by Lunine and Stevenson (1982). It is also reasonable to assume that the temperature profile has a form that can be approximated by $T \propto R^{-1}$, where R is the distance from the center of Jupiter. This functional form is suggested by numerical simulations of the solar nebula (Cameron, 1978).

A radial temperature profile

$$T = 3600(R_J/R)K, \tag{6}$$

then follows if the nebula mass out to ~ 30 R_J is ~ 0.1 M_J (Lunine and Stevenson, 1982). The condensation sequence for this profile is described in Prinn and Fegley (1982). This profile predicts that if the satellites formed in regions close to their present orbits, then Io should form from partially hydrated rock, Europa from hydrated rock, and Ganymede and Callisto should incorporate the solar abundance of water ice. This is consistent with the observed properties of these satellites. The requirement that the nebula be somewhat more massive than a minimal nebula is no problem since some condensed matter may be accreted into Jupiter or be ejected back into the solar nebula. Lunine and Stevenson also find that the observed

surficial differences between Ganymede and Callisto can be explained in this nebula model because Ganymede possessed a primordial water ocean by virtue of forming in a hotter part of the nebula, whereas Callisto may have formed as an unmelted mixture of ice and rock.

The nebula calculation outlined here is intended to be only suggestive and does not replace a detailed computer simulation, but it does indicate how the satellite properties can be related to the thermal regime of the forming planet. A similar calculation for Saturn, assuming $T \approx 11500 \, \rho^{0.44}$ for the initial adiabat, predicts $T \sim 60 \, K$ at Titan's orbit and thus indicates that the Saturnian satellites can incorporate the more volatile (CH_4, NH_3) ices as clathrates, unlike the icy Galilean satellite. This explains why Titan has an atmosphere, whereas Ganymede and Callisto do not. The incorporation of ammonia clathrate into the small Saturnian satellites may also explain why they appear to have undergone partial resurfacing, because partial melting by radiogenic heating to form an NH_3–H_2O magma can result in igneous activity. However, the lack of a clear density trend in the Saturnian satellite system (Smith et al., 1981) precludes as simple a model as the one describing the Galilean satellites. The choice of initial adiabats used in this discussion is considered further in the next section.

GASEOUS COLLAPSE CALCULATIONS

I turn now from inferences based on observations to purely theoretical calculations on formation scenarios. Of course, it would be nice to relate these scenarios to the observations, but it will become apparent that the relationships are often tenuous.

The pioneering calculations on the early evolution of Jupiter were by Grossman et al. (1972), subsequently refined by Graboske et al. (1975) and applied to Saturn (Pollack et al., 1977). More recent calculations (Grossman et al., 1980) incorporate the effects of a dense core and consider both Jupiter and Saturn. All these calculations are essentially equivalent to the Hayashi phase of stellar evolution calculations. The initial state, assumed non-rotating for simplicity, is an adiabatic body with a radius between one and two orders of magnitude larger than the present radius. The choice of initial radius is unimportant but the assumption of adiabaticity at all times is important. In these calculations, the luminosity is very high (almost 1% of the present solar luminosity) for the first 10^4–10^5 years, while the radius decreases by a large amount and the central temperature rises to a peak value ~50,000 K. The center then becomes degenerate (it is this which prevents Jupiter from collapsing still further and becoming a star). Subsequently, during the degenerate cooling phase, the luminosity decays roughly as $t^{-4/3}$ (where t is the elapsed time) and the contraction is much slower. This latter part of the evolution is necessarily governed by equation (2). Although all these calculations may reliably represent the degenerate cooling phase, they cannot be expected to provide accurate information on the first 10^5–10^8 years of evolution because of the artificiality of an initially adiabatic, homologously contracting state. The absence of rotation also precludes the formation of a disk, from which the satellites form. Pollack and Reynolds (1974) attempt to model the formation of the Galilean satellites by using the results of Graboske et al. (1975) but this must be regarded as an unsatisfactory procedure because the properties of a disk are very different from the environment of a hypothetical, spherically symmetric protoplanet. (For example, the disk is necessarily optically thick and the timescales for orbital evolution by gas drag are much smaller than the Kelvin cooling time.) Even in the optically thick model for a protoplanetary nebula discussed by Cameron and Pollack (1976), the cooling times were taken from Graboske et al. (1975); this is an invalid procedure.

Bodenheimer and co-workers have carried out a number of calculations incorporating hydrodynamic effects (Bodenheimer, 1974; Bodenheimer, 1977; Bodenheimer et al., 1980). These are essentially Jeans collapse calculations, starting from enormously distended isothermal states (initial radius is several thousand times the final radius). The evolution can be characterized by three stages: (i) an early, cool phase in which the hydrogen is molecular; (ii) a hydrodynamic collapse initiated by the dissociation of H_2 and ending when temperatures reach ~1×10^4 K and (iii) a final slow contraction and cooling which merges with the degenerate cooling phase described by equation (2). Although these calculations avoid the artificiality of the assumed adiabatic, homologous contraction of Grossman and co-workers, the initial state is still implausible because simple Jeans collapse cannot occur in the presence of the forming Sun. Nevertheless, many of the features in these calculations, such as the possibility of initiating collapse by the dissociation of H_2, are likely to be present in more realistic models.

Bodenheimer (1977) also considered the effects of rotation and found a tendency towards forming a binary system rather than a circumplanetary disk. This may be a consequence of omitting from the calculations the effect of eddy viscosity (caused by convection or mechanical mixing by accretion). The most important consequence of Bodenheimer's calculations is that they imply lower peak temperatures and luminosities than the adiabatic, homologous collapse calculations. Peak central temperatures of Jupiter and Saturn are about twice the present central temperatures; this is consistent with the choice of adiabats in my earlier discussion of the satellites. Existing published calculations of Bodenheimer and co-workers have not been carried out far enough into the onset of degeneracy to establish precisely how the models merge with the degenerate cooling phase predicted by equation (2).

Gaseous collapse calculations may be applicable to scenarios such as that developed by Cameron (1978), in which the solar nebula fragments into numerous giant, gaseous protoplanets. There are two fundamental and unresolved problems that models of this kind have yet to overcome: How do you explain the cores of the giant planets and the similarity of their masses? How do you explain the terrestrial planet systematics? The giant, gaseous protoplanets are expected to evolve in such a way as to cause rain-out of iron and silicates to the center (Slattery et al., 1980) but unless there is a continual supply of new material from the solar nebula, the resulting core is much smaller than the present day core. It is difficult to envisage how this scenario would lead to similar, large core masses for the giant planets especially since later incorporations of silicate and iron would be soluble in the deeper regions and could not rain out. The theory also predicts giant, gaseous protoplanets in the inner solar system. Even if the gaseous envelopes of such protoplanets could be eliminated (e.g. by evaporation or tidal stripping), no explanation of the present atmospheric rare gas systematics, especially [36]Ar, is apparent in such a model.

These troublesome difficulties lead to consideration of a very different model in which the cores form first and hydrodynamic collapse of the gaseous components follows.

NUCLEATED INSTABILITIES

Perri and Cameron (1974) considered the stability of a dense core surrounded by an adiabatic, cosmic composition envelope which merges uni-

formly with the temperatures and pressures of the solar nebula. They found that an instability occurs only for rather large (50–100 $M_{\oplus}$) cores. However, their assumption of adiabaticity is incorrect: it is straightforward to show that the net outward energy transport by radiation alone at a few optical depths in their models is greater than the present solar luminosity!

More realistic models have recently been constructed by Mizuno (1980). In these models, the luminosity is limited to the energy available from accretion and the outer regions of the gaseous envelope are more nearly isothermal than adiabatic. A striking consequence of this is that the thermal structure deep within the protoplanet is insensitive to the outer boundary conditions imposed by the solar nebula. This is profoundly different from adiabatic models, which require that the specific entropy of the interior is necessarily equal to that of the region of the solar nebula in which the protoplanet is immersed. As a consequence of the insensitivity to outer boundary conditions, Mizuno finds that hydrodynamic collapse occurs for essentially the same core mass regardless of where the core is placed in the solar nebula. This may explain the similarity of core masses for the giant planets.

In view of the potential importance of this result, it is useful to develop an approximate analytic theory that reproduces it. This approximate theory may also expose potential difficulties with the model. Consider a core of mass M_c surrounded by a gaseous envelope in which $p(r)$, $\rho(r)$ and $T(r)$ are the pressure, density and temperature at radial position r. Hydrostatic equilibrium requires that

$$\frac{1}{\rho}\frac{dp}{dr} = -\frac{GM_r}{r^2}, \qquad (7)$$

where M_r is the mass within radius r. If radiation transport dominates, then

$$\frac{16\sigma T^3}{3K\rho} \cdot \frac{dT}{dr} = -\frac{L}{4\pi r^2}, \qquad (8)$$

where K is the Rosseland mean opacity (assumed constant) and L is the luminosity at r. Dividing equation (8) and (7) and assuming both T, P are much larger than their outer boundary values gives

$$T^4 \simeq \frac{3KL}{16\pi\sigma GM} P. \qquad (9)$$

This is known by astrophysicists as the "radiative zero" solution. Substituting the ideal gas equation of state $p = (\rho/\mu)k_B T$, where μ is the molecular weight and k_B is Boltzmann's constant, yields.

$$T \simeq \frac{\mu G M_t}{4 k_B r} \qquad (10)$$

$$\rho \simeq \frac{64\pi\sigma}{3KL}\left(\frac{\mu G M_t}{4 k_B}\right)^4 \cdot \frac{1}{r^3}, \qquad (11)$$

where M_r is approximated by the total mass M_t and L is assumed to be constant. The temperature profile of equation (10) is convectively stable. The mass of the envelope, $M_{env} \equiv M_t - M_c$, is given by

$$M_{env} = \int_{R_c}^{R_0} 4\pi\rho(r) r^2 \, dr$$
$$= \frac{64\pi\sigma}{3KL}\left(\frac{\mu G M_t}{4 k_B}\right)^4 \cdot 4\pi \ln\left(\frac{R_0}{R_c}\right), \qquad (12)$$

where R_c is the core radius and R_0 is the radius of the photosphere. (Since it enters in a logarithmic term, the calculation is insensitive to R_0.) If the luminosity is derived from the accretion of rock/ice planetisimals, then

$$L \cong \frac{G M_c}{R_c} \cdot \frac{d M_c}{d t}. \qquad (13)$$

Thus, $L \propto M_c^{2/3}$ if the accretion rate is constant. Since $M_t = M_c + M_{env}$, it follows from equation (12) that

$$M_t \cong M_c + \left(\frac{\alpha}{K}\right)\frac{M_t^4}{M_c^{2/3}}, \qquad (14)$$

where weak (logarithmic) terms are incorporated in the parameter α, which is treated as a constant with respect to M_c or M_t but depends strongly on μ ($\alpha \propto \mu^4$) and is roughly proportional to the accretion timescale.

It is clear that equation (14) has no real solution for M_t if M_c is too large. The largest possible value of M_c which is consistent with hydrostatic equilibrium is found by setting $dM_c/dM_t = 0$, the solution to which is

$$M_t = \left(\frac{K}{4\alpha}\right)^{1/3} M_c^{2/9}$$

$$M_c \simeq 0.38\left(\frac{K}{\alpha}\right)^{3/4}. \qquad (15)$$

For the particular parameter choices of Mizuno

(1980), $M_c/M_\oplus \simeq 20\,K^{3/7}$ where K is in units of $cm^2\,g^{-1}$. The simple model is compared with Mizuno's calculation in Fig. 2. The physical interpretation is this: if you attempt to construct a protoplanet with a core mass larger than the critical value given by equation (15) then hydrostatic equilibrium cannot be satisfied. There is an unbalanced downward acceleration and hydrodynamic collapse ensues. The required core mass is predicted to be insensitive to outer boundary conditions, mainly because of the weak variation of $\ln R_0$.

Although this model provides an attractive explanation for the similarity of giant planet core masses, it does require a surprisingly large opacity $K \sim 1\,cm^2\,g^{-1}$ to reproduce the correct core mass. This opacity is not achievable by the constituent gases but could be provided by a cosmic abundance of ice and rock in the form of μm or sub-μm grains. If most of the condensibles have already been incorporated in large planetesimals, then the envelope could not be this opaque. However, incoming planetesimals could disseminate part of their constituents by ablation and break-up as they penetrate the gaseous envelope on their way to the core. This raises another difficulty which is not evident in Mizuno's calculations but reveals itself in the simple model: the
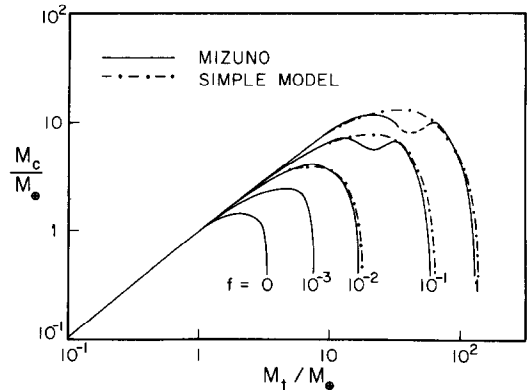

FIG. 2. CORE MASS (M_c) AS A FUNCTION OF TOTAL MASS (M_t) FOR A ROCK CORE SURROUNDED BY A GAS ENVELOPE IN HYDROSTATIC EQUILIBRIUM AND IMMERSED IN THE SOLAR NEBULA.
The solid curves are from Mizuno (1980) and the dashed lines are the simple model described in the text. The parameter f is a measure of opacity (essentially K in units of $cm^2\,g^{-1}$). In each case, there is a maximum possible value of M_c. For larger values, no hydrostatic equilibrium is possible and hydrodynamic collapse occurs.

critical mass scales as $\mu^{-12/7}$. Dissemination of incoming comet-like planetesimals can substantially increase μ before the predicted $\sim 10\,M_\oplus$ core is reached. Collapse may then be determined more by the requirement that the envelope be hot enough to accommodate evaporated ices. This requirement is satisfied for a core mass $\sim 1\,M_\oplus$. More calculations are needed before the nucleated collapse models can be said to be well understood. Perhaps numerous proto-giant planets are formed and then coalesce to form the present giant planets.

Two other major problems confront the nucleated instability model. The first concerns the determination of final mass: Why does Jupiter grow to $\sim 320\,M_\oplus$ whereas Saturn grows to $\sim 95\,M_\oplus$? (The total mass at the onset of instability in Mizuno's models is the same for each and about equal to the present masses of Uranus and Neptune.) It is too simplistic to say that these masses represent the total reservoirs of gaseous material available, since then one cannot logically explain the similarity of core masses! Something else is involved (e.g. preferential loss of gas from outer regions of the solar nebula). The other major problem in the nucleation model is that giant planet formation is delayed by the formation time of the cores. During this period ($\sim 10^7$ years), there should have been sufficient time to make the terrestrial planets. However, the small mass of Mars and the existence of the asteroid belt are usually explained by the prior formation of Jupiter. Furthermore, the noble gas systematics of the terrestrial planets are very difficult (if not impossible) to explain by accretion in the presence of a gaseous nebula (Wetherill, 1981). These considerations point towards forming the giant planets before the terrestrial planets.

CONCLUDING REMARKS

No current model of giant planet formation appears to satisfy all the constraints presented here. Models involving gaseous instabilities of the solar nebula fail to explain the presence of primordial rock/ice cores of similar mass in each of the giant planets and also encounter difficulties in the terrestrial planet region. Models involving collapse of a gaseous envelope onto a core of rock and ice may overcome these problems but are difficult to reconcile with the indications that the giant planets formed before the terrestrial planets. Either class of model appears capable of providing a disk from which the satellite systems could form.

Although the gaseous instabilities are "unavoidable" in the context of the particular accretion disk models envisaged by Cameron (1978) it is possible to conceive of satisfactory solar nebula models in which they would not occur. On the other hand, the evidence for the cores of the giant planets is incontrovertible (although there can still be uncertainty regarding their primordial nature). If one adopts the view that the cores are indeed the "seeds" of giant planet formation then the problems to be addressed are these.

(1) How did the giant planets form before the terrestrial planets?

(2) How do you avoid the problem that the nucleation may be too "easy"? (The required core may be much smaller than $\sim 10\,M_\oplus$.)

(3) How do you explain the final wide range of masses of the giant planets?

(4) In what ways can the model be tested?

I offer the following suggestions, merely to indicate that these problems may be tractable.

(1) Condensation, sedimentation and accretion may have begun much earlier in the outer parts of the solar nebula, where the temperatures were lower. Synchroneity of events in the outer and inner parts of the solar nebula is unlikely.

(2) Dissemination of icy planetesimals (ablating and breaking up as they plunge into the protoplanetary envelope) may create a convective state because of a density inversion. Convective mixing and thermal transport may modify the thermal structure of the envelope so as to delay the nucleated instability.

(3) Growth of a giant planet core is more rapid at Jupiter's orbit than further out. If gradual inflow or outflow were causing the gaseous nebula to decline in density during the epoch of giant planet formation, then Jupiter could have swept up more gas more quickly than the giant planets further out. Notice, however, that this would have to be reconciled with point 1, which tends to counteract this.

(4) A direct test of the nucleated instability model is difficult to envisage but indirect tests can be made by improving the current understanding of high pressure physics, especially the solubilities of rock and ice in metallic hydrogen. This knowledge, when integrated with detailed formation models, can lead to predicted structures and evolutions. These can be compared with the present planetary structures. It is conceivable that some of the present atmospheric abundances can be related to the formation process.

REFERENCES

Bodenheimer, P. (1974). Calculations of the early evolution of Jupiter. *Icarus* **23**, 319.

Bodenheimer, P. (1977). Calculations of the effects of angular momentum on the early evolution of Jupiter. *Icarus* **31**, 356.

Bodenheimer, P., Grossman, A. S., DeCampli, W. M., Marcy, G. and Pollack, J. B. (1980). Calculations of the evolution of the giant planets. *Icarus* **41**, 293.

Cameron, A. G. W. (1978) Physics of the primitive solar nebula and of giant gaseous protoplanets, in *Protostars and Planets* (Ed. T. Gehrels) pp. 453–487, University of Arizona Press, Tucson.

Cameron, A. G. W. and Pollack, J. B. (1976). On the origin of the solar system and of Jupiter and its satellites, in *Jupiter* (Ed. T. Gehrels) pp. 61–84, University of Arizona Press, Tucson.

Graboske, H. C., Pollack, J. B., Grossman, A. S. and Olness, R. J. (1975). The structure and evolution of Jupiter: the fluid contraction phase. *Astrophys. J.* **199**, 265.

Grossman, A. S., Graboske, H., Pollack, J., Reynolds, R. and Summers, A. (1972) An evolutionary calculation of Jupiter. *Phys. Earth. Planet. Interiors* **6**, 91.

Grossman, A. S., Pollack, J. B., Reynolds, R. T., Summers, A. L. and Graboske, H. C. (1980). The effect of dense cores on the structure and evolution of Jupiter and Saturn. *Icarus* **B42**, 358.

Hewitt, J. M., McKenzie, D. P. and Weiss, N. O. (1975). Dissipative heating in convective flows. *J. Fluid Mech.* **68**, 721.

Hubbard, W. B. (1980). Intrinsic luminosities of the Jovian planets. *Rev. Geophys. Space Phys.* **18**, 1.

Linden, P. F. and Shirtcliffe, T. G. L. (1978). The diffusive interface in double-diffusive convection. *J. Fluid Mech.* **87**, 417.

Lunine, J. I. and Stevenson, D. J. (1982). Formation of the Galilean satellites in a gaseous nebula, submitted to *Icarus*.

Mizuno, H. (1980). Formation of the giant planets. *Progr. theor. Phys.* **64**, 544.

Perri, F. and Cameron, A. G. W. (1974). Hydrodynamic instability of the solar nebula in the presence of a planetary core. *Icarus* **22**, 416.

Pollack, J. B., Grossman, A. S., Moore, R. and Graboske, H. C. (1977). A calculation of Saturn's gravitational contraction history. *Icarus* **30**, 111.

Pollack, J. B. and Reynolds, R. T. (1974). Implications of Jupiter's early contraction history for the composition of the Galilean satellites. *Icarus* **21**, 248.

Prinn, R. G. and Fegley, B., Jr. (1982). Kinetic inhibition of CO and N_2 reduction in circumplanetary nebulae: implications for satellite composition. *Astrophys. J.* (in press).

Slattery, W. L., DeCampli, W. M. and Cameron, A. G. W. (1980) Protoplanetary core formation by rain-out of minerals. *Moon and Planets* **23**, 381.

Smith, B. A. and 26 others (1981). Encounter with Saturn: Voyager 1 imaging science results. *Science, N.Y.* **212**, 163.

Stevenson, D. J. (1980). Saturn's luminosity and magnetism. *Science, N.Y.* **208**, 746.

Stevenson, D. J. (1982). Interiors of the giant planets. *A. Rev. Earth Planet. Sci.* (in press).

Stevenson, D. J. and Fishbein, E. (1981). The behavior of water in the giant planets. *Lunar planet. Sci. Abstracts* **10**, 1040.

Stevenson, D. J. and Salpeter, E. E. (1977). The phase diagram and transport properties of hydrogen–helium fluid planets. *Astrophys. J. Suppl.* **35**, 221.

Turner, J. S. (1973). *Buoyancy Effects in Fluids*, Cambridge University Press, Cambridge.

Wetherill, G. W. (1981). Solar wind origin of ^{36}Ar on Venus. *Icarus* **46**, 70.