The Astrophysical Journal, 536:L101–L104, 2000 June 20

POSSIBLE RAPID GAS GIANT PLANET FORMATION IN THE SOLAR NEBULA AND OTHER PROTOPLANETARY DISKS

Alan P. Boss
Department of Terrestrial Magnetism, Carnegie Institution of Washington, 5241 Broad Branch Road, NW, Washington, DC 20015-1305; boss@dtm.ciw.edu
Received 2000 March 27; accepted 2000 May 5; published 2000 June 13

ABSTRACT

Gas giant planets have been detected in orbit around an increasing number of nearby stars. Two theories have been advanced for the formation of such planets: core accretion and disk instability. Core accretion, the generally accepted mechanism, requires several million years or more to form a gas giant planet in a protoplanetary disk like the solar nebula. Disk instability, on the other hand, can form a gas giant protoplanet in a few hundred years. However, disk instability has previously been thought to be important only in relatively massive disks. New three-dimensional, "locally isothermal," hydrodynamical models without velocity damping show that a disk instability can form Jupiter-mass clumps, even in a disk with a mass (0.091 $M_\odot$ within 20 AU) low enough to be in the range inferred for the solar nebula. The clumps form with initially eccentric orbits, and their survival will depend on their ability to contract to higher densities before they can be tidally disrupted at successive periastrons. Because the disk mass in these models is comparable to that apparently required for the core accretion mechanism to operate, the models imply that disk instability could obviate the core accretion mechanism in the solar nebula and elsewhere.

Subject headings: accretion, accretion disks — hydrodynamics — planetary systems — solar system: formation

1. INTRODUCTION

Spectroscopic searches have discovered about three dozen very low mass companions to nearby solar-type stars (Marcy, Cochran, & Mayor 2000). These companions are widely assumed to be gas giant planets similar to Jupiter, an identification that has been strengthened by the discovery of the first transiting extrasolar planet (Charbonneau et al. 2000; Henry et al. 2000) and the resulting determination of a planetary radius and density close to that expected for a hot Jupiter (Burrows et al. 2000). However, most (about $\frac{3}{4}$) of these objects are more massive than Jupiter, with likely masses in the range 1–10 Jupiter masses ($M_{\rm Jup}$). The first confirmed system of such objects, around the star υ Andromedae, has a combined mass roughly 10 times that of our solar system's planets (Butler et al. 1999). Theories of gas giant planet formation derived to explain the formation of Jupiter may not be adequate to account for these considerably more massive objects.

Conventional wisdom suggests that Jupiter and Saturn formed by the core accretion mechanism (Mizuno 1980; Pollack 1984), which is a two-step process. First, collisional accumulation of icy and rocky planetesimals leads to the runaway growth (Lissauer 1987) of a $\sim$10 $M_\oplus$ solid core. Second, as this core grows, it acquires a growing atmosphere of nebular gas, which eventually becomes unstable to collapse, leading to a phase of rapid accretion of hydrogen and helium gas onto the protoplanet's envelope. The first step requires on the order of a half-million years, while the second step requires several million years or more according to the most recent models (Pollack et al. 1996). Core accretion became the favored mechanism for several reasons. The critical core mass required for gas accretion is not strongly dependent on orbital distance in the solar nebula (Mizuno 1980; Pollack 1984), offering an explanation for the then-inferred similarity in core masses for the giant planets (Stevenson 1982). By specifying the amount of nebular gas accreted during the rapid accretion phase, core accretion can lead to arbitrary ratios of core mass to envelope mass and hence to nonsolar bulk compositions. Finally, core accretion begins with the collisional accumulation of solids, the uncontested mechanism of terrestrial planet formation (Wetherill 1990, 1996).

However, core accretion is not without problems, some of which have only recently emerged. The perpetual problem has been the timescale, which is comparable to or longer than estimates of the lifetime of planet-forming disks. Such estimates range from a few million years or less in low-mass star–forming regions (Wolk & Walter 1996; Jayawardhana et al. 1999) to much less than one million years in high-mass star–forming regions (Bally et al. 1998), where most stars are born. If the disk gas has already been dissipated, a massive gaseous envelope cannot be accreted. Other problems are more recent. New models of the interior of Jupiter suggest a considerably smaller core mass or even no core at all (Guillot, Gautier, & Hubbard 1997). If the core masses are not large enough to initiate sustained gas accretion, the core accretion mechanism will fail. Furthermore, a 10 $M_\oplus$ core is expected to migrate inward toward the protosun in $\sim 10^4$ yr through gravitational interactions with the nebula, well before it can accrete a significant gaseous envelope (Goldreich & Tremaine 1980; Ward 1997a, 1997b; Tanaka & Ida 1999; Miyoshi et al. 1999). If a core manages to avoid this fate by accreting a sufficiently massive gaseous envelope, it will open a gap in the disk, slowing the further growth of the planet (Bryden et al. 1999).

The only alternative mechanism appears to be disk instability, where a gravitationally unstable disk fragments directly into self-gravitating clumps of gas and dust that can contract and become giant gaseous protoplanets. Disk instability was first proposed decades ago (Kuiper 1951; Cameron 1978), but was largely discarded because disk instability seemed to be unable to form large solid cores, as solids are expected to dissolve in the hot envelope rather than settle to the center of a gas giant planet (Slattery, DeCampli, & Cameron 1980; Stevenson 1982). However, a solid core could form by coagulation

TABLE 1
Effects of Increasing Spatial Resolution on Disk Instability Models

Model	Damping	N_ϕ	N_{Ylm}	ρ_{max}
64d	Yes	64	16	6.3×10^{-8}
64f	No	64	16	7.9×10^{-10}
128f	No	128	16	1.6×10^{-9}
256f	No	256	16	2.0×10^{-9}
256pf	No	256	32	1.6×10^{-8}
512pf	No	512	48	4.3×10^{-7}

of dust grains and their sedimentation to a protoplanet's center in $\sim 10^3$ yr (Boss 1997, 1998a), well before the protoplanet contracts to planetary densities and temperatures, which requires $\sim 10^5$ yr (Bodenheimer et al. 1980). A 1 M_{Jup} protoplanet of solar composition ($Z = 0.02$) could then quickly form a $\sim 6\ M_\oplus$ solid core, a value that lies in the middle of the presently inferred range for Jupiter (Guillot et al. 1997).

The disk instability process itself is quite fast, as it occurs on orbital timescales, so that clump formation and dust grain sedimentation proceed nearly simultaneously, on a timescale of $\sim 10^3$ yr. A disk instability is able to form multiple-M_{Jup} planets (Boss 1998a), because the more massive (and cooler) a disk is, the more likely it is to undergo the instability. Because the massive clumps form rapidly and only afterward open up disk gaps, there is no problem with growing to the large masses inferred for many extrasolar planets. Similarly, the disk instability mechanism avoids entirely the danger of significant orbital migration prior to reaching 1 M_{Jup}. Once a gap forms around a giant protoplanet formed by either mechanism, the protoplanet will be subject to orbital migration driven by the subsequent evolution of the disk, a likely means for explaining the short orbital periods of the hot Jupiters (Lin & Papaloizou 1986; Lin, Bodenheimer, & Richardson 1996; Trilling et al. 1998).

The disk instability mechanism also has problems. Gravitationally unstable disks are thought to evolve as a result of gravitational torques that transport mass inward and angular momentum outward and thereby avoid forming long-lived clumps (Cassen et al. 1981; Papaloizou & Savonije 1991; Laughlin & Bodenheimer 1994). A marginally unstable disk may then require a trigger to induce the instability, such as episodic mass accretion onto the disk (Boss 1997) or a close encounter with another star (Boffin et al. 1998). The instability may require a fairly massive disk (Cassen et al. 1981; Laughlin & Bodenheimer 1994; Boss 1998a), considerably more massive than the minimum mass necessary to make the solar system. As a result, disk instability may not be able to make gas giant planets as low in mass as Saturn or gas giant planets with nonsolar bulk compositions unless their gaseous envelopes are preferentially lost by overflow through their Roche lobes during inward orbital migration.

Here we demonstrate that disk instability can operate in a disk with a mass comparable to that inferred for the solar nebula and indeed seemingly required for core accretion to succeed, removing one of the major problems with disk instability.

2. HYDRODYNAMICAL MODELS

Previous three-dimensional hydrodynamics models (Boss 1997, 1998a) showed that the instability could proceed in a disk with a mass of 0.14 $M_\odot$ inside a radius of 10 AU; however, the total mass for a disk extending beyond 10 AU could be significantly higher. These models had a 1 $M_\odot$ central protostar, so that the ratio of the disk mass to the star mass was at least $M_d/M_s = 0.14$. Semianalytical studies of a specific eccentric disk instability mechanism suggested that instability could occur only in disks with $M_d/M_s > 0.19$ (Shu et al. 1990). Several numerical studies have found clumping to occur when $M_d/M_s \sim 1$ (Cassen et al. 1981; Laughlin & Bodenheimer 1994). One study of two-dimensional (thin) disks found clumping to occur in smoothed particle hydrodynamics calculations of disks with $M_d/M_s = 0.05$ and 0.1, but was unable to confirm the result with finite differences calculations; however, for $M_d/M_s \geq 0.2$, both numerical methods yielded qualitatively similar behavior (Nelson et al. 1998).

Estimates of a lower bound on the mass of the solar nebula fall in the range 0.01–0.1 $M_\odot$ (Weidenschilling 1977), so that the solar nebula must have had a ratio $M_d/M_s \geq 0.01$–0.1. The numerical models presented here have a disk mass of 0.091 $M_\odot$ within a radius of 20 AU and a 1 $M_\odot$ central protostar, so that $M_d/M_s = 0.091$.

The finite differences code used here solves the three-dimensional equations of hydrodynamics and the Poisson equation on a spherical coordinate grid. The code is the same as that previously employed (Boss 1997, 1998a; with the exception noted below) and has been shown to be accurate to second order in space and time (Boss & Myhill 1992). The number of grid points in each direction is $N_r = 101$, $N_\theta = 23$ in $\pi/2 \geq \theta \geq 0$, and $N_\phi = 64$, 128, 256, or 512 (Table 1). The radial grid is uniformly spaced between either 1 or 4 AU and 20 AU, with boundary conditions at both 1 (or 4) AU and 20 AU chosen to absorb radial velocity perturbations. The θ grid is compressed into the midplane to ensure adequate vertical resolution ($\Delta\theta = 0.3°$ at the midplane). The ϕ grid is uniform. The central protostar is allowed to wobble in response to the growth of nonaxisymmetry in the disk, thereby preserving the location of the center of mass of the star/disk system. The number of terms in the spherical harmonic expansion for the gravitational potential of the disk is varied from $N_{Ylm} = 16$ to 32 to 48.

The disk initially has a surface density profile with $\sigma \propto r^{-1/2}$–r^{-1} in the inner disk, steepening to $\sigma \propto r^{-3/2}$ in the outer disk, similar to that thought to be appropriate for the core accretion mechanism (Lissauer 1987). The surface density at 5 AU falls within the likely bounds for the solar nebula (Weidenschilling 1977) and is about 50% higher than in the standard core accretion model (Pollack et al. 1996). The initial disk density distribution is seeded with a mixture of nonaxisymmetric cos $(m\phi)$ density perturbations (m = 1, 2, 3, and 4 modes with amplitude $a_m = 0.01$) and with random noise at a somewhat lower amplitude.

The three-dimensional models start with the thermal structure of the corresponding axisymmetric disk as calculated by a two-dimensional radiative hydrodynamics code (Boss 1996). Because the axisymmetric model extended to only 10 AU, whereas the new models extend to 20 AU, the temperature profile from 10 to 20 AU was taken to be constant at 40 K. This appears to be a reasonable temperature for the outer solar nebula, based on a variety of cosmochemical constraints (Boss 1998b). With this temperature profile, the disk initially has a (Toomre 1964) $Q_{min} \sim 1.3$; marginally unstable disks have Q_{min}-values less than ~ 1.5 (Papaloizou & Savonije 1991; Nelson et al. 1998; Boss 1998a). The disk is assumed to evolve as a "locally isothermal" disk, meaning that the initial radial temperature profile is held fixed (Boss 1997, 1998a).

Previous models by Boss (1997, 1998a) damped the disk's translational velocities whenever $v_r > 0$ or $v_\theta < 0$ in order to maintain a stable inner disk without suffering a severe time step penalty. Marginally unstable disk models calculated by Pickett et al. (1998, 2000a), however, did not include this damping and did not lead to the formation of the long-lived clumps found by Boss (1997, 1998a). The present study includes models with and without velocity damping in order to learn the effects of this artifice with the present code (see Pickett et al. 2000b).

3. RESULTS

Table 1 summarizes the six models and the approximate maximum density encountered during the evolution. As the spatial resolution is increased, the maximum density encountered increases, as expected for calculations that are converging toward a solution involving clump formation.

Model 64d was calculated in much the same manner as previous models (Boss 1997, 1998a), with limited ϕ spatial resolution and velocity damping. After ∼3000 yr, the result was similar: two multiple-$M_{\rm Jup}$ clumps formed and orbited apparently stably on circular orbits. Model 64f was identical to model 64d except for having the velocity damping removed beyond 5 AU. Model 64f became nonaxisymmetric within ∼300 yr and still produced multiple-$M_{\rm Jup}$ clumps, but the clumps were on eccentric orbits and disappeared after at most a single revolution period. Thin filaments persisted throughout the evolution, and new clumps continued to form out of them and then disappear. Model 128f, with doubled ϕ resolution, behaved much like model 64f, but the clumps reached maximum densities about a factor of 2 higher. Evidently, velocity damping had the unintended effects of preserving the clumps indefinitely and of slowing the evolution, as has also been found by Pickett et al. (2000b). For the subsequent models, velocity damping was removed completely and the inner boundary was moved out to 4 AU.

Model 256f again doubled the ϕ resolution, but the maximum density achieved by the clumps increased only slightly, implying convergence in the sense of the hydrodynamical grid. However, because clump formation is driven by self-gravity, it is important to also increase the resolution of the Poisson solver. Model 256pf had doubled N_{Ylm} = 32 compared to model 256f and led to almost a factor of 10 increase in the maximum clump density. After this model had formed a few clumps, it was continued from that point with N_ϕ = 512 and N_{Ylm} = 48. As shown in Figure 1, this model, 512pf, led to an even higher maximum clump density, about 27 times higher than in model 256pf and about 7 times higher than that encountered in the initial velocity-damped model 64d. The clump shown in Figure 1 survived for two orbital periods (60 yr). As N_ϕ and N_{Ylm} are increased, the clumps become better defined and reach much higher densities but still disappear if the calculation is run indefinitely. This suggests that even much higher spatial resolution is needed to follow the clumps faithfully and that the formation of the clumps is not the result of insufficient spatial resolution.

At the time shown in Figure 1, the most prominent clump has a mass of 5.2 $M_{\rm Jup}$, compared to a Jeans mass of 0.02 $M_{\rm Jup}$ at the average density of the clump, showing how tightly gravitationally bound the clump is. The clump also has an average spherical radius of 0.2 AU, considerably smaller than the critical tidal radius of 1.2 AU for its mass, orbital distance, and the central protostar mass, implying stability to tidal disruption (Boss 1998a). The clump's free-fall time at its maximum density is about 0.1 yr, a small fraction of its orbital period of about 30 yr, implying that the clump should be able to contract toward protoplanetary densities if it were allowed to by the spatial resolution of the grid. The clump's orbital eccentricity is $e \sim 0.3$, suggesting that disk instability could lead to protoplanets on initially noncircular orbits.

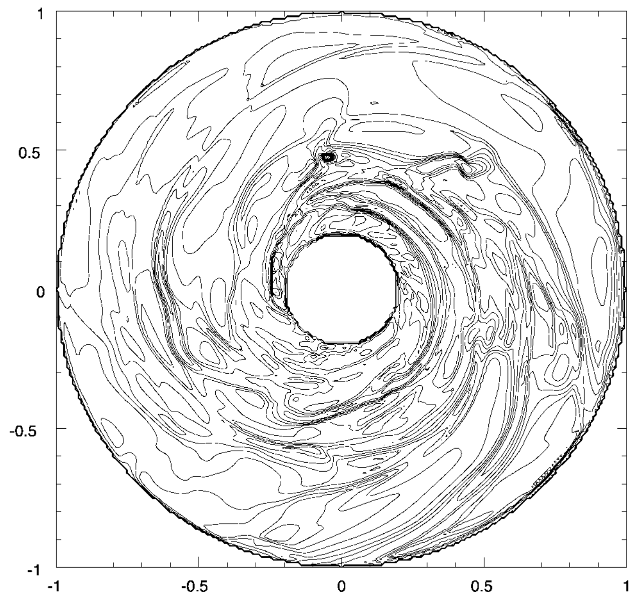

FIG. 1.—Midplane density contours for the highest spatial resolution model (512pf) at 374 yr. Contours represent factors of 2 change in density. Diameter of region shown is 40 AU; inner hole has a diameter of 8 AU. Filaments, clumps, and holes are evident. The maximum density of 4.3×10^{-7} g cm^{-3}, which is particularly well-defined, occurs in the clump near 12 o'clock, survives for two orbits, and is likely to contract toward planetary densities in a calculation with even higher spatial resolution.

4. CONCLUSIONS

The models show that strongly self-gravitating clumps can form in a marginally gravitationally unstable disk even without velocity damping. While the clumps appear to be capable of forming gas giant protoplanets, their subsequent orbital evolution and survival are highly uncertain. The models suggest that a phase of disk instability is likely to be a relatively chaotic environment for planet formation. If several clumps survive and reach planetary densities, the resulting gas giant planets may suffer mutual close encounters, resulting in even higher orbital eccentricities.

Disk instability leading to the formation of gas giant protoplanets thus seems to be possible in a disk with a mass (∼0.1 $M_\odot$) comparable to the upper end of the range inferred for the solar nebula and other protoplanetary disks (e.g., Beckwith et al. 1990). The surface density at 5 AU in the model is comparable to that favored in standard core accretion models (Pollack et al. 1996). However, in order to avoid loss of the gaseous envelope by a hydrodynamic blowoff, core accretion may require that the planet be embedded in a disk with a density at 5 AU of ∼10^{-9} g cm^{-3} (Wuchterl, Guillot, & Lissauer 2000), a density that is at least 5 times higher than that of the disk instability models presented here and comparable to previous disk instability models (Boss 1997, 1998a).

Because the timescale for disk instability is roughly 10^3 times shorter than that for core accretion, clearly if it can occur, a disk instability will circumvent core accretion in the solar nebula or in disks orbiting around other protostars. However, it would be premature to attempt to decide which of these mechanisms is superior at forming gas giant planets, based on our present knowledge. For example, the models presented here assume "locally isothermal" thermodynamics, an assumption that errs strongly in favor of clumping and merits considerable further scrutiny (Boss 1997; Pickett et al. 1998, 2000a; Nelson, Benz, & Ruzmaikina 2000). The survival of the clumps found here is likely to depend to a large extent on a proper treatment of their thermodynamical evolution, which will be the subject of a future paper.

I thank Patrick Cassen, Stephen Kortenkamp, and Brian Pickett for discussions and the referee for improvements to the manuscript. Supported in part by the NASA Planetary Geology and Geophysics Program under grant NAG5-3873. Calculations were performed on the Carnegie Alpha Cluster, which is supported in part by NSF MRI grant AST 99-76645.

REFERENCES

Bally, J., Testi, L., Sargent, A., & Carlstrom, J. 1998, AJ, 116, 854
Beckwith, S. V. W., Sargent, A. I., Chini, R. S., & Güsten, R. 1990, AJ, 99, 924
Bodenheimer, P., Grossman, A. S., DeCampli, W. M., Marcy, G., & Pollack, J. B. 1980, Icarus, 41, 293
Boffin, H. M. J., Watkins, S. J., Bhattal, A. S., Francis, N., & Whitworth, A. P. 1998, MNRAS, 300, 1189
Boss, A. P. 1996, ApJ, 469, 906
———. 1997, Science, 276, 1836
———. 1998a, ApJ, 503, 923
———. 1998b, Annu. Rev. Earth Planet. Sci., 26, 53
Boss, A. P., & Myhill, E. A. 1992, ApJS, 83, 311
Bryden, G., Chen, X., Lin, D. N. C., Nelson, R. P., & Papaloizou, J. C. B. 1999, ApJ, 514, 344
Burrows, A., Guillot, T., Hubbard, W. B., Marley, M. S., Saumon, D., Lunine, J. I., & Sudarsky, D. 2000, ApJ, 534, L97
Butler, R. P., Marcy, G. W., Fischer, D. A., Brown, T. W., Contos, A. R., Korzennik, S. G., Nisenson, P., & Noyes, R. W. 1999, ApJ, 526, 916
Cameron, A. G. W. 1978, Moon Planets, 18, 5
Cassen, P. M., Smith, B. F., Miller, R. H., & Reynolds, R. T. 1981, Icarus, 48, 377
Charbonneau, D., Brown, T. M., Latham, D. W., & Mayor, M. 2000, ApJ, 529, L45
Goldreich, P., & Tremaine, S. 1980, ApJ, 241, 425
Guillot, T., Gautier, D., & Hubbard, W. B. 1997, Icarus, 130, 534
Henry, G. W., Marcy, G. W., Butler, R. P., & Vogt, S. S. 2000, ApJ, 529, L41
Jayawardhana, R., Hartmann, L., Fazio, G., Fisher, R. S., Telesco, C. M., & Piña, R. K. 1999, ApJ, 521, L129
Kuiper, G. P. 1951, Proc. Natl. Acad. Sci., 37, 1
Laughlin, G., & Bodenheimer, P. 1994, ApJ, 436, 335
Lin, D. N. C., Bodenheimer, P., & Richardson, D. C. 1996, Nature, 380, 606
Lin, D. N. C., & Papaloizou, J. C. 1986, ApJ, 309, 846
Lissauer, J. J. 1987, Icarus, 69, 249
Marcy, G. W., Cochran, W. D., & Mayor, M. 2000, in Protostars & Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: Univ. Arizona Press), 1285
Miyoshi, K., Takeuchi, T., Tanaka, H., & Ida, S. 1999, ApJ, 516, 451
Mizuno, H. 1980, Prog. Theor. Phys., 64, 544
Nelson, A. F., Benz, W., Adams, F. C., & Arnett, W. D. 1998, ApJ, 502, 342
Nelson, A. F., Benz, W., & Ruzmaikina, T. V. 2000, ApJ, 529, 357
Papaloizou, J. C., & Savonije, G. J. 1991, MNRAS, 248, 353
Pickett, B. K., Cassen, P., Durisen, R. H., & Link, R. 1998, ApJ, 504, 468
———. 2000a, ApJ, 529, 1034
Pickett, B. K., Durisen, R. H., Cassen, P., & Meija, A. C. 2000b, in Poster Proc., IAU Symp. 200, Birth and Evolution of Binary Stars, ed. R. Mathieu & H. Zinnecker (Potsdam: Astrophys. Inst.), 190
Pollack, J. B. 1984, ARA&A, 22, 389
Pollack, J. B., Hubickyj, O., Bodenheimer, P., Lissauer, J. J., Podolak, M., & Greenzweig, Y. 1996, Icarus, 124, 62
Shu, F. H., Tremaine, S., Adams, F. C., & Ruden, S. P. 1990, ApJ, 358, 495
Slattery, W. L., DeCampli, W. M., & Cameron, A. G. W. 1980, Moon Planets, 23, 381
Stevenson, D. J. 1982, Planet. Space Sci., 30, 755
Tanaka, H., & Ida, S. 1999, Icarus, 139, 350
Toomre, A. 1964, ApJ, 139, 1217
Trilling, D. E., Benz, W., Guillot, T., Lunine, J. I., Hubbard, W. B., & Burrows, A. 1998, ApJ, 500, 428
Ward, W. R. 1997a, ApJ, 482, L211
———. 1997b, Icarus, 126, 261
Weidenschilling, S. J. 1977, Ap&SS, 51, 153
Wetherill, G. W. 1990, Annu. Rev. Earth Planet. Sci., 18, 205
——. 1996, Icarus, 119, 219
Wolk, S. J., & Walter, F. M. 1996, AJ, 111, 2066
Wuchterl, G., Guillot, T., & Lissauer, J. J. 2000, in Protostars & Planets IV, ed. V. Mannings, A. P. Boss, & S. S. Russell (Tucson: Univ. Arizona Press), 1081