

Available online at www.sciencedirect.com



Icarus 170 (2004) 492-507

ICARUS

www.elsevier.com/locate/icarus

Planetary migration in a planetesimal disk: why did Neptune stop at 30 AU?

Rodney S. Gomes^{a,*}, Alessandro Morbidelli^b, Harold F. Levison^c

^a GEA/OV/UFRJ and ON/MCT, Ladeira do Pedro Antônio, 43—Centro 20.080-090 Rio de Janeiro, RJ, Brazil ^b Observatoire de la Côte d'Azur, Boulevard de l'Observatoire, BP. 4229, 06304 Nice, France

^c Southwest Research Institute, Department of Space Sciences, 1050 Walnut St. Suite 429, Boulder, CO 80302, USA

Received 1 July 2003; revised 25 March 2004

Available online 28 May 2004

Abstract

We study planetary migration in a gas-free disk of planetesimals. In the case of our Solar System we show that Neptune could have had either a damped migration, limited to a few AUs, or a forced migration up to the disk's edge, depending on the disk's mass density. We also study the possibility of runaway migration of isolated planets in very massive disk, which might be relevant for extra-solar systems. We investigate the problem of the mass depletion of the Kuiper belt in the light of planetary migration and conclude that the belt lost its pristine mass well before that Neptune reached its current position. Therefore, Neptune effectively hit the outer edge of the proto-planetary disk. We also investigate the dynamics of massive planetary embryos embedded in the planetesimal disk. We conclude that the elimination of Earth-mass or Mars-mass embryos originally placed outside the initial location of Neptune also requires the existence of a disk edge near 30 AU.

© 2004 Elsevier Inc. All rights reserved.

Keywords: Kuiper belt; Resonance; Solar System formation

1. Introduction

Planet migration in forming planetary systems occurs in two stages. The first one happens due to the interaction of the planet with the gaseous disk (Ward, 1997; Masset, 2001). After the gas disk dissipates, the energy and angular momentum exchange between remaining planetesimals and the planets induce the second stage of planetary migration. This phenomenon was first brought to light by Fernandez and Ip (1984).

It is now believed that planetary migration substantially sculpted the Kuiper belt, generating most of the features that are now observed. Malhotra (1993) first showed that the resonant, eccentric orbit of Pluto can be the result of the 2:3 resonance sweeping through the proto-planetary disk during Neptune's migration. Similarly, the same scenario explains the existence of a significant fraction of Kuiper belt bodies in the major motion resonances with Neptune 3:4,

* Corresponding author. Fax: +(55)-21-2263-0685. *E-mail address:* rodney@ov.ufrj.br (R.S. Gomes). 2:3, 3:5, 1:2, and their wide range of orbital eccentricities (Malhotra, 1995). Gomes (2003) showed that the origin of the so called 'hot classical Kuiper belt' (a population of non-resonant bodies with inclinations larger than 4°) can also be explained as a result of Neptune's migration, which allowed a small portion of the scattered disk population to be trapped on stable orbits with small/moderate eccentricities. More recently Levison and Morbidelli (2003) proposed that Neptune's migration also generated the 'cold classical Kuiper belt' (the population of non-resonant bodies with inclinations smaller than 4° ; Brown, 2001): the members of this population would have been transported to their current location from a much smaller heliocentric distance through a mechanism that invokes temporary trapping into the 1:2 mean motion resonance.

The properties of the Kuiper belt are not the only indications of planetary migration. Levison and Stewart (2001) showed that the in situ formation of Uranus and Neptune is unlikely, suggesting that these planets formed much closer to Jupiter and Saturn, where the growth timescales were dramatically shorter (Thommes et al., 2003). Thommes et al. (1999) proposed a radical different view, in which Uranus

^{0019-1035/\$ –} see front matter $\,$ 2004 Elsevier Inc. All rights reserved. doi:10.1016/j.icarus.2004.03.011

and Neptune formed between Jupiter and Saturn and were scattered outwards, where the interactions with the disk of planetesimals damped their eccentricities and inclinations.

Despite the importance of planetary migration, not much work has been done up to now to study the migration process per se. After the pioneering work of Fernandez and Ip (1984), Hahn and Malhotra (1999) tried to better characterize planetary migration with a series of direct numerical integrations. In their work, the planets, initially in a more compact configuration, were embedded in a planetesimal disk with total mass ranging from 10 to 200 Earth masses (M_{\oplus}) , and with a surface density decaying as the inverse of the heliocentric distance r. Because of computational limitations, the authors were forced to simulate the disk with only 1000 objects, which exerted a gravitational influence on the planets but not among themselves. The authors found that a 50 M_{\oplus} disk could bring Neptune from its initial position, postulated at 23 AU, to its quasi-final position at 30 AU in 50 Myr, and therefore concluded that this was the most likely mass of the planetesimal disk after planetary formation. An important point observed in Hahn and Malhotra (1999) is that migration proceeded in a non-adiabatic way, so that no resonance trapping of the planetesimals was observed. The authors conjectured that, if the disk were composed of a larger number of smaller planetesimals, Neptune's migration would be smoother and consequently the resonance trapping phenomenon would occur. This, they argued, could also slow Neptune's outward motion because the resonant particles would effectively increase Neptune's inertial mass (as they need to be moved together with the planet).

Gomes (2003) simulated Neptune's migration using a disk of 10,000 massive planetesimals. As expected, he observed a much smoother migration than in Hahn and Malhotra (1999), with many resonant captures. However, despite the captures, with a disk similar to that of Hahn and Malhotra (60 Earth masses between 20 and 45 AU with a $r^{-1.5}$ surface density profile), Neptune migrated to 45 AU in 1.4×10^8 yr. The fact that this result was so different from the one by Hahn and Malhotra shows the necessity of a deeper understanding of the phenomenon of planetary migration, which is precisely the goal of the present paper.

A detailed study of the general migration process would require the exploration of a huge parameter space and thus is beyond our current technical ability. Thus, we limit ourselves to explore the cases that, we believe, might be the most instructive to understand the primordial evolution of our Solar System.

We start in Section 2 with a simple analytical model that stresses the exponential character of the migration process. This will be useful to interpret the results of the numerical simulations presented in the next sections. In Section 3, we discuss migration in large-mass disks. In Section 4, we consider the case of low-mass disks and discuss how the resolution of the simulation (number of massive planetesimals used to model the disk) affects the simulation results. Section 5 addresses the issue of the depletion of the primordial mass of the Kuiper belt and its effects on Neptune's migration. We rule out the possibility that the belt was depleted by some dynamical mechanism that moved most Kuiper belt bodies to Neptune-crossing orbit. We also argue that the Kuiper belt could not have lost its mass by collisional grinding after that the planet reached 30 AU. We therefore conclude that Neptune stopped at its current location because it encountered an effective edge of the massive proto-planetary disk. Then, in Section 6, we discuss, in detail, Neptune's migration in truncated disks and deduce the range of plausible disk masses and sizes that are compatible with the current position of Neptune. We also investigate the implications for the Thommes et al. (1999) scenario. Section 7 discusses what would have been the dynamical evolution of planetary embryos, if they existed in the disk beyond Neptune's primordial position. Our conclusions will be recollected in Section 8. The appendix reports the details on the integration methods that we have used.

2. A simple analytic insight in the migration process

In this section we develop a back-of-the-envelope analytic 'theory' for migration in planetesimal disks. Our goal is to present an intuitive, easy to understand toy model, intended to be a guide for interpreting the range of behaviors observed in our numerical simulations. We refer the reader to Ida et al. (2000b) for a more developed analytic theory.

The consequences of the encounter between two bodies in orbit around the Sun can be effectively computed in most of the cases using an impulse approximation (Opik, 1976). In this approximation the effect of the encounter is an instantaneous rotation of the orbital velocity vectors of the two bodies, computed using the well-known Rutherford twobody scattering formulae. Using this approach, it is easy to compute (Valsecchi and Manara, 1997) that on average (that is averaged on all impact parameters and relative orientations) the planetesimals that cause an outward migration to a planet on a circular orbit are those whose z-component of the angular momentum $H = \sqrt{a(1-e^2)}\cos i$ is larger than that of the planet, H_p . The opposite is true for the planetesimals with $H < H_p$. In these formulae *a*, *e*, and *i* are the semimajor axis, eccentricity and inclination of the planetesimal. This is due to the fact that, when encountering the planet, the particles with $H > H_p$ have on average a velocity component in the direction tangential to the planet's motion that is larger than the orbital velocity of the planet. Thus they accelerate the planet. The opposite is true of the particles with $H < H_p$. This result applies also if the planet has a moderate eccentricity.

The direction of migration of the planet is therefore determined by the relative populations of planet-crossing planetesimals with $H > H_p$ and $H < H_p$. This may be different from case to case. Some general trends, however, can be outlined. For instance, in the case of two planets, the inner planet partially depletes the population of planetesimals with



Fig. 1. Semi-major axis and eccentricity of the planets (filled dots) and of the planetesimals (points) at time $t = 2.3 \times 10^6$, from the simulation with a 200 Earth masses disk initially in the 20–50 AU region presented in Section 3. The solid lines define the limits for planetary-crossing orbits, while the dotted lines show where $H = H_p$ for zero inclination orbits.

H smaller than the angular momentum of the outer planet, so that the latter tends to migrate outwards. Similarly, the outer planet partially depletes the population of planetesimals with H larger than the angular momentum of the inner planet, so that the latter tends to migrate inwards.

In our Solar System, migration should have had a general trend, with Jupiter moving inward and Neptune moving outward. Figure 1 shows an example of semi-major axis vs. eccentricity distribution of the planets and the planetesimals during the migration process. For each planet, the solid curves show the boundaries of the planet-crossing regions and the dotted curves correspond to the condition $H = H_p$ for i = 0. The overlapping of the Neptune-crossing and Uranus-crossing regions implies a gradual depletion of the objects with $H < H_{Neptune}$ relative to those with H > H_{Neptune} . The consequence of this imbalance induces the outward migration of Neptune. The same reasoning can be applied to the other planets except for Jupiter. Jupiter is so massive that it rapidly ejects to the interstellar space most of the planetesimals that come close to its orbit (or sends a small fraction into the Oort cloud), so that it must move inwards (this mechanism has been proposed for the origin of the hot Jupiters in extra-solar system by Murray et al., 1998). Notice however that the situation might have been temporarily different if the planets encountered discontinuities in the surface density distribution of the planetesimal disk, such as gaps, edges, or density clumps (possibly caused by the migration itself), which, in some cases, could cause reversals in the direction of migration (see Section 3 for examples).

For a better understanding of the numerical simulations presented next, we first develop an analytic toy model of the migration process.

During migration, the fractional rate of change of the planet's semi-major axis, da'_p/dt , where $da' = da/a_p$, is proportional to: (1) the ratio of amount of material in orbits that cross the orbit of the planet, M(t), to the mass of the planet, M_p , (2) a function k of the distribution of those orbits (for example the distribution of H described above), and (3) the timescale between close encounters between small particles and the planet, which in turn is proportional to 1/P, where P is the orbital period of the planet ($P = 2\pi a_p^{3/2}$). Therefore

$$\frac{\mathrm{d}a_p}{\mathrm{d}t} = \frac{k}{2\pi} \frac{M(t)}{M_p} \frac{1}{\sqrt{a_p}}.$$
(1)

Note that most of the dynamics of the system is hidden in the parameter k. The evolution of M(t) can be approximated by the equation

$$\dot{M}(t) = -M(t)/\tau + 2\pi a_p |\dot{a}_p|\sigma(a_p), \qquad (2)$$

where the first term in the r.h.s. represents the decay of the planetesimal population due to the planetesimal's finite dynamical lifetime, and the second term stands for the planetesimals that, because of the change in the planet's position, enter for the first time the region where they can be scattered by the planet. In (2) $\sigma(a_p)$ is the surface density of the 'virgin' (i.e., not yet scattered) planetesimal disk at heliocentric distance a_p . Substituting (1) into (2) we get

$$\dot{M}(t) = \left(-\tau^{-1} + |k|\sqrt{a_p}\sigma(a_p)/M_p\right)M(t).$$
(3)

Let us assume for simplicity that the term $\alpha = -\tau^{-1} + \tau^{-1}$ $|k| \sqrt{a_p} \sigma(a_p) / M_p$ does not significantly change with time (an approximation valid for small migrations, but which evidently looses its validity when the migration covers a macroscopic range). Then, (3) becomes an exponential equation with solution $M(t) = M(0) \exp(\alpha t)$. If α is negative, then M(t) decays exponentially to 0 and the planet (from Eq. (1)) stops migrating. In this case, the loss of planetesimals due to their finite dynamical lifetime is not compensated by the acquisition of new planetesimals in the scattering region, because the migration speed is too slow. Therefore, the planet runs 'out of fuel.' We call this migration mode *damped* migration. Conversely, if α is positive, M(t) grows exponentially and the planet exponentially accelerates (Ida et al., 2000b). We call this migration mode forced migration. In this case the acquisition of new planetesimals due to the migration exceeds the losses due to the finite dynamical lifetime, and the migration is self-sustained.

The description of migration through Eqs. (1) and (2) is necessarily crude. In reality, the migration can pass from damped to a forced mode and vice-versa, as the surface density σ , the decay time τ and the relative planetesimal distribution k change with the planet's location a_p and planetary migration rate \dot{a}_p . The changes of τ and k along the migration cannot be estimated, a priori. Also, if \dot{a}_p becomes large

enough, disk particles can cross the planet's zone of influence in a timescale short compared to τ and thus leave the zone from the opposite edge. This effect introduces a new negative term in the r.h.s. of (2) that also cannot be evaluated a priori, because it depends on the details on the interactions between the particles and the planet. Planetesimals can also be trapped in mean motion resonances, which effectively increases the inertial mass of the planet, thus causing a decrease of k. On the other hand, planetesimals exterior to the planet's zone of influence may be dynamically excited to planet-crossing orbits by resonances. This increases the delivery of fresh mass to the planet compared to the term $2\pi a_p |\dot{a}_p| \sigma(a_p)$ in (2). Moreover, the relative orbital distribution of the planetesimals in the planet-crossing zone may change during the evolution, causing a change in k. Finally, the width of the planet-crossing region changes linearly with a_p , also modifying the r.h.s. of (2).

Therefore, this system of equations cannot be effectively used to *simulate* the migration process. Indeed, we are not aware of any theory on planetary migration in planetesimal disks that can substitute for numerical simulations. However, our toy model shows the intrinsic exponential nature of the migration process, and therefore will be very useful to *interpret* the results of the direct simulations of the migration process that will be presented in the next sections.

3. Migration in large mass disks

We present two simulations of giant planet migration due to the presence of a massive disk. In both cases, the initial semi-major axes of Jupiter, Saturn, Uranus, and Neptune are 5.4, 8.7, 13.8, and 18.1 AU, respectively and their initial eccentricities are 0.

In the first case, the disk extends from 12.5 to 45 AU. Following the idea that the disk should be strongly depleted in the planetary region due to the accretion of the planets, we assume a disk mass of $4.6M_{\oplus}$ inside 20 AU, and of 117.5 M_{\oplus} outside 20 AU, with a surface density decaying as $r^{-1.5}$ in each sub-region. The disk is simulated using 6123 equal-mass particles. In the second case, the disk extends up to 50 AU, and contains $203M_{\oplus}$ outside 20 AU, with a surface density decaying as r^{-1} , while the mass inside 20 AU is again equal to $4.6M_{\oplus}$. It is simulated using 10,190 equal mass particles. These surface density profiles are those typically assumed for the protoplanetary disk (Hayashi, 1981; Hahn and Malhotra, 1999). In these and all other simulations presented in this paper, the disk particles responded to the planets, but not to each other. We are aware that this approximation, imposed by the necessity to keep the computing time within reasonable limits, introduces some artifacts. The frequencies of secular precession of the disk particle orbits are not correct, which misplaces the location of the secular resonances with the planets. Also, collective effects are not reproduced, which suppresses a torque that would subtract angular momentum from the planets (Goldreich and

Fig. 2. Evolution of the semi-major axes of the four giant planets due to a planetesimal disk of $\sim 200M_{\oplus}$ initially between 20 and 50 AU (top) and $\sim 120M_{\oplus}$ initially between 20 and 45 AU (bottom). A low mass disk of $4.6M_{\oplus}$ is assumed in both cases in the 12.5–20 AU range.

Tremaine, 1980). We think that these artifacts do not have severe consequences. Unlike mean motion resonances (see below), secular resonances do not play an important role in planetary migration, except for possibly providing additional distant sources of planetesimals to the planet-crossing region. Collective effects become unimportant as soon as the planetesimal disk becomes moderately excited (Ward and Hahn, 1998a, 1998b).

Figure 2 shows the migration of the planets for the two planetesimal disks defined above. For the reasons explained in Section 2, Saturn, Uranus, and Neptune on average migrate outward, while Jupiter migrates inward. Neptune undergoes forced migration because these disks are massive. Consequently, the planet eventually migrates to the edge of the disk (and in fact goes slightly beyond it), and it can come to a rest only when the disk has been mostly depleted (which occurs at about 2×10^7 yr). The fact that Neptune's real position is at 30 AU obviously rules out the idea that a similar extended massive disk was present in our Solar System at early times.

However, other planetary systems might have had in the past disks of comparable mass and (even larger) radial extent, and therefore migration may have brought planets to large distances from their parent stars. Such planets have been postulated to explain features observed in the disks around β Pictoris (Wahhaj et al., 2003), Vega and ϵ Eridani (Ozernoy et al., 2000). If the observational evidence for their existence is substantiated, we believe that forced migration in a massive planetesimal disk might be a valid explanation of their origin. However, the migration process as described here requires that the planet is much less massive than the disk and is incapable of ejecting most planetesimals to hy-



perbolic orbit; thus it only applies to planets like Uranus and Neptune rather then Jupiter and Saturn.

Another important result shown in Fig. 2 is that Neptune's migration is not monotonic. In the case of the high-mass disk (top panel), Neptune reaches ~ 50 AU in less than 4 Myr, and then comes back to within 30 AU almost equally fast. Similar episodes of acceleration and return (although less pronounced) are also visible in the low-mass disk case, at ~ 4 and ~ 10 Myr. This behavior is due to a self-sustaining migration process, described in Ida et al. (2000b), that we call *runaway migration*.

Under normal conditions, the planetesimals in Neptunecrossing orbit that have $H < H_{Neptune}$ are depleted by Uranus. Therefore, there is never a large number of H < H_{Neptune} particles that could drive Neptune inward, so that Neptune's outer migration is irreversible. But when Neptune migrates fast or gets far from Uranus, it can get in a mode where it does not scatter objects into Uranus-crossing orbits. These objects are therefore left behind in an excited disk as Neptune moves forward (compare Fig. 3 to Fig. 1). However, planet's outward migration continues as long as the planetesimals in the Neptune-crossing zone with $H > H_{\text{Neptune}}$ dominate over those with $H < H_{\text{Neptune}}$. The two populations do not rapidly equilibrate because of the migration itself, which continuously supplies new planetesimals with $H > H_{Neptune}$ to the Neptune-crossing region (Ida et al., 2000b). However, when Neptune reaches the edge of the disk and thus the number of objects with $H > H_{Neptune}$ drops, (Fig. 3B), the remaining objects interior to Neptune pull the planet inward. Thus, Neptune reverses direction and starts a runaway inward migration. The same argument described above applies, so that this migration ends only when the region of the disk partially depleted by Uranus is encountered again.

To demonstrate that Uranus has no role in the runaway migration of Neptune or in its reversal, we perform the following experiment. At t = 3 Myr we remove all planets except Neptune and extend the disk's outer edge to 60 AU following the original surface density distribution. Then we continue the integration with only Neptune and the planetesimal disk. Figure 4 compares Neptune's evolution in the new simulation to its previous one. As expected, the two evolutions show essentially the same behavior before ~ 3.3 Myr. However, in the new integration, Neptune continues its migration until it reaches the new edge. In a third integration, we extend the disk up to 80 AU. Again, Neptune continues its migration up to the new edge. This series of simulations show that, once started, runaway migration proceeds without the help of the other planets, and that hitting the disk's edge causes the migration to be reversed.

However, Neptune's behavior suddenly changes when we extend the disk further. In the integrations shown by the top curves in the figure, we extend the disk up to 200 AU. Neptune migrates much further than in the previous cases, but surprisingly, it but does not reach the new edge. In all integrations that we have made (7 in total, 3 shown in Fig. 4),



Fig. 3. Semi-major axis and eccentricity distribution of the planetesimals at $t = 3.1 \times 10^6$ (panel A) and $t = 3.2 \times 10^6$ yr (panel B) for the simulation presented in the top panel of Fig. 2. The lines define the boundaries of the planet-crossing regions.

Neptune reverses its migration at $\sim 110-120$ AU. We note that Neptune is not more likely to eject planetesimals from the Solar System when it is further from the Sun because, for a particle encountering the planet, the probability to be ejected to hyperbolic orbit depends exclusively on the Tisserand parameter, and the latter is independent of the semi major axis units. Therefore the reversion of Neptune's migration requires a more subtle explanation.

In order to understand the reversal in Neptune's migration, we first must understand why the entire migration process seems to proceed with a quasi-periodic alternation of accelerations and slow-downs (or even stops, see Fig. 4). Referring back to Eq. (1), we believe that this is due to combination of two effects as the migration proceeds: a slow



Fig. 4. Neptune's migration in the simulation presented in the top panel of Fig. 2 is shown here on a magnified timescale. Other curves show Neptune's migration in new simulations in which Jupiter, Saturn, and Uranus are discarded at 3 Myr, and the disk is extended beyond 50 AU. The similarity between the previous and the new evolutions up to 3.3 Myr demonstrates that the other planets play an inessential role in Neptune's runaway migration. In the simulations with the outer edge of the disk at 60 and 80 AU, the planet migrates up to the edge, and then reverses migration. In the simulations with the disk's edge at 200 AU, the inversion occurs well before the edge is reached.

decrease in k and an increase in M. As we described in Section 2, a planet migrates outward if more of the disk particles encountering it have $H > H_p$ than have $H < H_p$ (which implies that k > 0), or it migrates inward if the opposite is true (and k < 0). Figure 5 shows the density of Neptune-crossing particles at three time-steps that correspond to the beginning, the middle, and the end of a fast migration episode. The figure clearly shows that, initially, particles cluster in the $H > H_p$ region, but progressively move towards the $H < H_p$ region as the migration proceeds. This shift in the distribution of the particles happens because, when the planet migrates sufficiently fast, the timescale for encountering the planet becomes comparable to or longer than that for passing through the planet-crossing region due to the migration of the planet itself. So, in a coordinate system that moves with the planet, most particles simply drift through a significant fraction of this region before suffering an encounter. Thus, when the planet sees the particle for the first time its H is significantly smaller than the value that characterized the particle when it first became planet-crosser, and can even be smaller than H_p . The net result is that k slowly decreases with time.

At the same time, we find that the amount of mass in the planet-crossing region (*M*) increases with time. The value of *M* changes as $2\pi a_p \sigma(a) \Delta$, where Δ is the width of the planet-crossing region (remember that in this case there is no dynamical depletion of the planet-crossing particles). Because $\Delta \propto a_p$ and in this problem the surface density of the disk is proportional to the inverse heliocentric distance, $M \propto a_p$.

So, *M* is increasing while *k* is decreasing. Since $\dot{a}_p \propto kM\sqrt{a_p}$, if *k* decreases with time more slowly than $1/\sqrt{a_p}$,



Fig. 5. Greyscale-coded density of planetesimals in Neptune-crossing regions at three different times corresponding to the beginning of a fast migration episode (left panel), the middle (middle panel), and the end of it (right panel). The migration is then reversed. The semi-major axis is expressed in units of the current semi-major axis of the planet, in order to highlight the differences among the planetesimals distributions. The continuous light grey curves mark the borders of Neptune-crossing region, and the dashed curve the condition $H = H_p$ for i = 0.

the magnitude of \dot{a}_p actually increases with time. This happens until *k* becomes equal to 0, at which point migration abruptly stops. From Fig. 4, it seems that this phenomenon becomes somewhat more pronounced as the planet gets further from the Sun. In fact, the timescale for encounters grows as $a_p^{3/2}$, i.e., faster than the width of the planet-crossing region (proportional to a_p), so that it becomes easier to shift the distribution of the planet-crossing particles (as it happens in Fig. 5) and reduce *k*.

Now that we have understood why Neptune's migration repeatedly stops, we can now discuss the migration reversal seen at $a \sim 120$ AU. At every stopping episode, Neptune finds itself in an unstable situation. If the planet stays at rest for a long enough time, the excitation of the outer cold disk due to secular and resonant perturbations eventually brings new material into the planet-crossing region with $H > H_p$, so that the planet starts migrating outward again. This happens every time that a new acceleration of the migration is produced in Fig. 4. But, if the excited disk interior to Neptune (which is made up of $H < H_p$ particles) slightly overpowers the particles from the outer disk, the planet starts to migrate inward. This is enough to trigger a runaway inward migration, because the planet finds a massive excited disk inside its orbit, ready to refill the planet-crossing region, while the cold outer disk is left behind. We have not been able to identify any dynamical reason for why, in some cases, Neptune sometimes reverses direction. Thus, we believe it is a matter of chance. If so, this whole effect may be



Fig. 6. Neptune's semi-major axis evolution for planetesimals disks with several surface density distributions and total masses. Each disk was modeled using 10,000 particles. In the cases labeled 40, 45, and 50 M_{\oplus} , the disk has a surface density decaying as r^{-1} and total masses of 40, 45, 50 M_{\oplus} between 18 and 50 AU, respectively. In the cases labeled $r^{-1.5}$ and r^{-2} the disk has a total mass of $50M_{\oplus}$ and a surface density decaying as $r^{-1.5}$ and r^{-2} , respectively.

the result of the fact that our simulations contain a relatively small number of massive bodies compared to the real early Solar System. Perhaps an ideal system with a nearly infinite number of planetesimals with infinitesimal mass would behave differently. We will address this issue again in future work.

The possibility that Neptune may have had a period of inward migration if it were embedded in a massive disk suggests a new mechanism for the excitation of the classical Kuiper belt: Neptune might have crossed the belt and then returned to 30 AU, dynamically exciting the Kuiper belt in its wake. Unfortunately, this scenario cannot work. In the simulations that we performed of this process (see Fig.1), after its inward migration, Neptune always reverses its migration once again and eventually reaches the original outer edge of the disk. Therefore Neptune could not have stopped at 30 AU, but would have reached a final position beyond the Kuiper belt. Moreover, if Neptune had ended its travels immediately after a period of inward migration, the population of the Plutinos would probably not have survived. In fact, during a period of inward migration the particles in exterior mean motion resonances experience a decrease in eccentricity, until they are eventually released from the resonance.

4. Migration in low mass disks

We now investigate the migration process for disk masses smaller than $50M_{\oplus}$. In all our simulations, the initial locations of Jupiter, Saturn, Uranus, and Neptune are 5.45, 8.7, 15.5, and 17.8 AU, respectively. The disk extends from 18 to 50 AU and has a surface density variation as r^{-1} . It is simulated using effectively 10,000 equal mass particles, although we employ a computational trick to decrease the amount of CPU time the runs require (see Appendix A).

Figure 6 shows Neptune's migration for disk masses equal to 40, 45, and 50 M_{\oplus} . The first two cases are examples of damped migration (see Section 2). Neptune's outward motion rapidly slows down, and the planet reaches, after 10^9 yr, a quasi-asymptotic distance that is well within the outer edge of the disk. The part of the disk outside a few AUs beyond Neptune preserves its original mass, while the part within this distance is completely depleted. These results are qualitatively equivalent to those obtained by Hahn and Malhotra (1999) for disks of 10 and 50 M_{\oplus} . In our case, Neptune stops at ~ 24 and ~ 26 AU, respectively, but it started more than 5 AU closer to the Sun than in Hahn and Malhotra's simulations.

When we increase the disk mass to $50M_{\oplus}$, we observe a change of behavior. Neptune's outward migration first slows down, then stays approximately linear between 100 and 600 Myr, and finally accelerates towards the disk's edge. This evolution suggests that the surface density of this disk approximately corresponds to the critical one that separates dumped migration from forced migration (see Section 2). We believe that the acceleration of Neptune's migration seen after 600 Myr is due to the following. In these simulations, as Neptune migrates outward some of the particles in the external disk become trapped in Neptune's mean motion resonances (Malhotra, 1993, 1995). These particles are dragged outward with Neptune's migration, but their eccentricities are pumped up during this process. The resonant particles effect migration because they effectively increase Neptune's inertial mass. If the migration rate is slow enough so that the changes that the particles see are adiabatic, they stay in the resonance until they reach some critical eccentricity at which point they are released. During adiabatic migration, the number of particles in the resonances is roughly constant as long as the resonance is still in the disk.

In this run, Neptune accelerates as its 1:2 mean motion resonance moves out of the disk. This is most likely due to the fact that the number of objects in the resonance drops because objects in the resonance are leaving as their eccentricity grows, but new particles are not being captured. As the resonance is depleted, the effective inertial mass of Neptune decreases, which, in turn, we believe, causes Neptune to accelerate. We also believe that this acceleration is then amplified by the fact that Neptune starts moving quickly enough so that its migration becomes non-adiabatic. So, the resonant capture efficiency drops for all the resonances, and thus the total number of objects in resonances decreases. This belief is supported by the fact that we observed a decrease in the number of objects in Neptune's 2:3 resonance that starts soon after the drop in the 1:2. We should point out that this last effect could have happened even if the 1:2 did not hit the end of the disk.

The fact that our result for a $50M_{\oplus}$ disk is qualitatively different from that of Hahn and Malhotra should not be a concern. This mass is close to the threshold for the transition from damped to forced migration. As is usually the case when a physical system is studied close to a threshold, small quantitative differences in the simulations can lead to qualitatively different results. The major difference between our simulations and Hahn and Malhotra's is the different number of particles used to represent the planetesimal disk (10,000 particle sin our case, 1000 in Hahn and Malhotra's case). To illustrate how the number of particles matters, we have redone the simulations using only 1000 particles in the disks, as in Hahn and Malhotra (1999). The results are shown in Fig. 7. Two simulations are done for each disk's mass with different, but equivalent, initial conditions for the disk's particles. In all cases, we notice a large variability of the results. In particular, for the disks with 40 and 45 M_{\oplus} , in one simulation Neptune stops its migration inside 30 AU, as in Fig. 6, but in the other case it migrates towards the edge of the disk. In the 50 M_{\oplus} case both simulations lead Neptune to 46 AU, but the evolution paths are quite different. Unfortunately, we cannot prove the same problem does not exists in our 10,000 particle runs illustrated in Fig. 6. However, the fact that in Fig. 6 the final position of Neptune shows a regular progression with the disk mass, makes us think that stochasticity of Neptune's migration should be much less prominent.



Fig. 7. Neptune's semi-major axis evolution in a pair of simulations with disks of 40 (bottom), 45 (middle), and 50 M_{\oplus} (top). The surface density of the disks decays as r^{-1} and each disk is modeled with only 1000 particles. Due to the low resolution of the disk model, Neptune's migration results highly stochastic and unpredictable.

We can understand this stochastic behavior of Neptune's migration in low resolution disks on the basis of the analytic insight of Section 2. If the disk's surface density is close to the critical value that separates damped migration from forced migration, the evolution becomes very sensitive to the density fluctuations. If the disk is modeled by a small number of massive particles, the density fluctuations are more pronounced and stochastic, while if the disk is modeled with a larger number of smaller particles, the density fluctuations are more effectively averaged out in space and time. In particular, the encounters of Neptune with planetesimals with too large a mass inhibit the resonance trapping process (as pointed out by Hahn and Malhotra), thus changing the orbital distribution of the planetesimals that drive Neptune's migration. Also, in the case of a smaller number of more massive planetesimals, Neptune's eccentricity is larger on average (roughly 0.015 compared to ~ 0.005 for the 10,000 particles runs), which changes the dynamics in three ways: (1) Neptune's-crossing region is larger so it is easier for particles to become Neptune-crossing, (2) Neptune's resonances become stronger so that the external disk is more easily excited, and (3) Neptune can more easily change a particle's Tisserand parameter so that dynamical evolution occurs more quickly.

Figure 6 also shows two examples of migration obtained with disks of $50M_{\oplus}$, but surface densities decaying as $r^{-1.5}$ and r^{-2} . Here, 10,000 particles are used to model the disk, in both cases. A steeper surface density implies more mass in the inner part of the disk and less mass in the outer part. Therefore the migration starts faster than in the case of the r^{-1} surface density but, when the planet reaches the outer part of the disk, the locally low surface density puts it in a damped migration mode. These two effects (a faster initial migration and a slower final migration) combine in such a way that the resulting total migration of the planet becomes smaller as the surface density distribution of the disk gets steeper (for equal total disk masses). In fact, the total angular momentum of the disk decreases with steeper surface density profiles Thus, for a given total disk mass there must be a steep enough surface density distribution that makes Neptune stop at 30 AU. For a disk of $50M_{\oplus}$ between 20 and 50 AU, simulations show that the required density profile is approximately r^{-4} . The problem with this steep profile is that, if true, the mass originally in the 40-50 AU region would be $1.5M_{\oplus}$ —an order of magnitude smaller than that required to grow, in situ, the large Kuiper belt objects that are observed (Stern and Colwell, 1997a; Kenyon and Luu, 1999).

5. Neptune's position and the mass depletion of the Kuiper belt

If we assume that the primordial Kuiper belt must have contained at least $\sim 10 M_{\oplus}$ between 40–50 AU in order to grow objects currently observed, the results of the previous section suggest a scenario similar to the one proposed by Hahn and Malhotra (1999)-the surface density of the disk was shallow (exponent ~ -1), the disk contained $\sim 45 M_{\oplus}$ of material between 20 and 50 AU and Neptune started at ~ 22 AU. The initial location of Neptune, which is 4 AU further outward than in the simulations of Fig. 6, was chosen so that it would stop migrating at \sim 30 AU after a damped migration. Notice that this scenario is in conflict with the conclusions of Levison and Stewart (2001), in which Uranus and Neptune had to form significantly closer than 20 AU from the Sun. Our understanding of planetary formation is not yet secure enough to confidently rule out that that Neptune formed beyond 20 AU. However, the scenario sketched above has another problem. If Neptune stopped at 30 AU because its migration was damped, then the disk beyond \sim 35 AU would have preserved its original large mass. The current mass of the region, inferred from the observations (Jewitt et al., 1996; Chiang and Brown, 1999; Trujillo et al., 2001; Gladman et al., 2001), is now less than $0.1M_{\oplus}$. Could the Kuiper belt loose most of its mass without substantially modifying Neptune's final location?

Two general mechanisms have been proposed for the mass depletion of the Kuiper belt: (1) the dynamical excitation of most bodies to the Neptune-crossing orbits after which they were ejected, and (2) the collisional comminution of most of the mass of the Kuiper belt into dust.

The dynamical depletion mechanism was first proposed by Morbidelli and Valsecchi (1997) and Petit et al. (1999). In their scenario, a planetary embryo, with mass comparable to that of Mars or of the Earth, was scattered by Neptune onto



Fig. 8. A self-consistent simulation of the Petit et al. (1999) scenario for the excitation and dynamical depletion of the Kuiper belt. Neptune is originally assumed at ~ 23 AU and an Earth-mass embryo at ~ 27 AU. Both planets are embedded in a $30M_{\oplus}$ disk, extending from 10 to 50 AU. The pair of black curves show the evolution of Neptune's perihelion and aphelion distance, while the grey curves refer to the embryo. Notice that the embryo is never scattered by Neptune. It migrates through the disk faster than Neptune until the disk's outer edge. Neptune interacts with most of the mass of the disk, thanks to the dynamical excitation of the latter due to the presence of the embryo. Therefore, it migrates much further that it would if the embryo were not present, and reaches a final position well beyond 30 AU.

an elliptic orbit that crossed the Kuiper belt for $\sim 10^8$ yr. The repeated passage of the embryo through the Kuiper belt excited the eccentricities of the Kuiper belt bodies. The vast majority of these became Neptune crossers and were subsequently dynamically removed. In the Petit et al. (1999) integrations that studied this scenario, however, the Kuiper belt bodies were treated as massless test particles, and therefore their ejection did not alter the position of Neptune.

Thus, we have re-done a Petit et al.-like simulation in the framework of a more self-consistent model, where Jupiter, Saturn, Uranus and Neptune are initially at 5.40, 8.78, 16.25, and 23.14 AU, respectively, an Earth-mass embryo is on a circular orbit at 27.15 AU, and the disk has $30M_{\oplus}$ between 10 and 50 AU, with a surface density profile decaying as r^{-1} . The result is shown in Fig. 8. The discussion of the dynamical evolution of the embryo is postponed to Section 7. Here we focus on the result that, despite the low mass of the disk (only $7.5M_{\oplus}$ between 40 and 50 AU), at the end of the integration Neptune has migrated well beyond 30 AU. Indeed, in this simulation, 15% of the disk particles are still in the system at the end, so we do not get enough dynamical depletion. In order to determine how much Neptune would migrate if we removed all of the particles, we continued this simulation and placed another Earth-mass embryos outside of Neptune (at 44 AU). In this integration Neptune reaches 37 AU after 1.5 Gyr, while 4% of the disk particles are still in the system.

In the above simulations, Neptune migrates further than it normally would without the embryos because the embryos dynamically excite the disk exterior to Neptune and feed this extra mass to it. Thus, Neptune interacts not only with the

portion of the disk in its local neighborhood, but with the entire mass of the disk at the same time. Therefore, a $30M_{\oplus}$ disk-which in absence of the embryo would allow Neptune to migrate only few AUs in a negative feedback modebrings Neptune well beyond its current position. We have done other numerical experiments with a set up equivalent to that of the simulation reported in Fig. 8, but different disk masses. If one requires that Neptune stopped at 30 AU, the disk in the 10–50 AU range should contain only $\sim 15 M_{\oplus}$ of planetesimals, the exact values depending on the initial location of the planet. This disk mass and density profile, however, would imply that only $3.75M_{\oplus}$ of material originally existed in the Kuiper belt between 40 and 50 AU, which is far less than the mass required $(10-30M_{\oplus})$ by the models of accretion of Kuiper belt bodies (Stern and Colwell, 1997a; Kenyon and Luu, 1999). Therefore, we believe that we can rule out the Petit et al. scenario for dynamical depletion of the Kuiper belt.

Although we have only studied the Petit et al. scenario, we believe that our results can be applied to *all* dynamical depletion scenarios. This is because Neptune's response to the mass once it leaves the Kuiper belt is unlikely to depend on whether Kuiper belt objects are excited to Neptunecrossing orbits by a planetary embryo or by some other mechanism, such as the primordial secular resonance sweeping (Nagasawa and Ida, 2000). Our results simply imply that Neptune never encountered the missing planetesimals of the Kuiper belt. Thus, the only type of dynamical deletion mechanism that could work is one in which the Kuiper belt objects were kicked directly to hyperbolic or Jupiter-crossing orbit and consequently were eliminated without interacting with Neptune. Only the passage of a star through the Kuiper belt is capable in principle of such an extreme excitation (Ida et al., 2000b; Kobayashi and Ida, 2001). However, a simple model (in which Neptune is at already at 30 AU, the Kuiper belt objects are fully formed, and a passing star causes the mass depletion and the dynamical excitation of the Kuiper belt) can probably be ruled out because it is unlikely to produce a Kuiper belt that is consistent with other observational constraints (Morbidelli and Brown, 2004).

An alternative mechanism for removing the mass from the early Kuiper belt is the *collisional grinding* scenario proposed by Stern and Colwell (1997b) and Davis and Farinella (1997, 1998). A massive Kuiper belt with large eccentricities and inclinations would undergo a very intense collisional activity. Consequently, most of the mass originally incorporated in bodies smaller than 50–100 km in size could be comminuted into dust, and then evacuated by radiation pressure and Poynting–Robertson drag. This would cause a substantial mass depletion, provided that the bodies larger than 50 km (which cannot be efficiently destroyed by collisions) initially represented only a small fraction of the total mass.

The collisional grinding of the Kuiper belt would not have been without consequences for Neptune's migration. The calculations of collisional grinding thus far performed have been sophisticated particle-in-a-box simulations that handle the evolving size-distribution of the Kuiper belt in a narrow annulus about the Sun by populating an array of mass bins. They then follow how the number of objects in each bin changes due to collisions. The smallest bin in the array is called the 'dust' bin and any mass put in this bin is subsequently ignored. But, what really happens to this dust? This depends on the size, shape, and composition of the particles, which may be all characterized by a single parameter-the ratio of the strength of radiation pressure to the strength of gravity, β (see Gustafson, 1994, for a review). Particles with $\beta \gtrsim 0.5$ are blown directly from the Solar System and thus do not interact with any other object. However, $\beta \lesssim 0.5$ dust particles spiral inward due to Poynting-Robertson (P-R) drag. For Kuiper belt dust, this means it will encounter Neptune (see Liou et al., 1996; Liou and Zook, 1999; Moro-Martín and Malhotra, 2003). If the dust created during the collisional grinding of the Kuiper belt has a size-distribution similar to that of the zodiacal cloud $(N(>R) \propto R^b)$, where R is particle radius and b = -1.2; Grogan et al., 2001) then much more mass will be found in the large particles than in the small particles. Thus, most of the dust (by mass) generated will spiral inward. Indeed, if we assume a particle with $R = 1 \ \mu m$ has $\beta = 0.5$ (Gustafson, 1994) and that only particles with $R < 100 \,\mu m$ respond to radiative forces (this is a very conservative upper limit, but choosing a larger one strengthens our case), then \gg 99% of the mass in dust will spiral toward the Sun and encounter Neptune. If we assume that the particles follow a collisional cascade size distribution (b = -2.5) this fraction is $\sim 91\%$. In either case the role of blow-off is negligible and almost all the dust will spiral inward.

So, the natural question is: How would Neptune respond to tens of Earth-masses of dust sailing by during the collisional grinding phase of the disk? Would it migrate outward as if it were interacting with larger particles? If so, collisional grinding could not be responsible for the mass depletion because Neptune would have migrated too far, as our earlier simulations have shown.

However, the response of Neptune is not obvious because, in part, as the dust particles migrate inward they get temporarily trapped in mean motion resonances (MMRs) with Neptune (Liou and Zook, 1999; Moro-Martín and Malhotra, 2003). In an MMR, the inward drift is halted because the energy loss due the radiation effects is balanced by the resonant interaction with the planet. The net result is that energy is extracted from the planet's orbit, so that the particles in the resonance try to drag Neptune in with them. This could be significant given that total mass of the dust generated during the collisional phase is comparable with the mass of Neptune. However, at the same time, these particles would have been slowly leaking out of the resonances and subsequently encountering Neptune. Like their larger brethren, this dust would have tried to push Neptune outward.

Thus, it is not clear exactly how Neptune will respond to the dust. We are currently studying this issue but it is very



Fig. 9. The location of the v_8 secular resonance as a function of the disk mass, assuming Neptune at 30 AU and the disk inner and outer boundaries at 35 and 50 AU. As the disk mass decreases, the v_8 secular resonance sweeps the disk. In the collisional grinding scenario, this phenomenon should have provided new material to the Neptune-crossing region and restart Neptune's migration.

unlikely that the two effects cancel out and Neptune's semi major axis remains approximately unchanged.

A way around this problem is that the dust is collisionally comminuted very quickly down to a size at which it is blown away by radiation pressure. In this case, it would not spend a significantly long time in Neptune-crossing orbit or in a mean motion resonance with the planet (Kenyon, private communication). However, even in this case collisional grinding would indirectly affect Neptune's orbit, due to the evolution of the v_8 secular resonance during the mass depletion phase. As the disk grinds down, the v_8 secular resonance most likely will begin to feed material to Neptune, which will then migrate. The v_8 secular resonance occurs when the periapse precession of a Kuiper belt object matches that of Neptune. This resonance is very powerful and any object in it suffers an increase in eccentricity until it can be removed from the Kuiper belt by a close encounter with Neptune (Holman and Wisdom, 1993; Duncan et al., 1995). It is currently at 40 AU. However, the presence of a massive disk (or annulus) affects the orbital precession frequencies of both Neptune and the disk particles. As the disk's mass grinds down the precession frequencies change. Consequently secular resonances move, potentially sweeping through the disk and exciting objects to Neptune-crossing orbits.

For example, assuming that, when Neptune reaches 30 AU, the disk has already been depleted inside 35 AU but is still massive in the 35–50 AU region, we have computed the location of the v_8 secular resonance as a function of the remaining disk's mass (Fig. 9) using the following semi-analytic model. The location of the secular resonance is simply the location where precession rate of the disk

particles is the same as that of the dominate frequency of Neptune. For the disk particles on low-inclination nearly circular orbits, the precession frequency can be determined using the epicyclic approximation (cf. Section 3.2 of Binney and Tremaine, 1987). In particular, it is very nearly equal to the difference between the radial oscillation frequency and the angular orbital frequency. The latter quantities can be computed from the first and second radial derivatives of the total gravitational potential produced by the massive Kuiper belt and the giant planets. For our model, we approximated the orbits of the planets by individual rings and the density distribution of the disk by a series of 5000 rings spread between 30 and 50 AU. We determined the first and second derivatives of the potential of the rings numerically. We calculate the dominate precession frequency of Neptune by developing a full secular theory of a system containing the four giant planets and seven fictitious planets with masses and semi-major axes chosen so that they approximate the disk.

As the figure shows, initially the v_8 resonance is at the inner edge of the disk. However, as the disk's mass decreases below $\sim 10M_{\oplus}$, the secular resonance starts sweeping through the disk. The resonance will begin to excite disk particles to Neptune-crossing orbits.¹ Because the disk still contains a lot of mass, about $0.5-1M_{\oplus}$ of material (assuming that the secular resonance is 2 AU wide) would start to have encounters with Neptune, forcing the planet to migrate outward. This, in turn, would move the resonance to a fresh location in the disk further from the Sun, which in turn would feed more particles to Neptune. In short, our guess is that an instability would be triggered, which would feed the remaining disk particles to Neptune and thus, as we showed for the dynamical depletion mechanisms above, Neptune would be driven beyond 30 AU.

In conclusion, we tend to exclude the possibility that collisional grinding depleted the mass of the Kuiper belt *after* that Neptune ended its damped migration at 30 AU. Of course, all the arguments discussed above can be circumvented if collisional grinding occurred during Neptune's migration, in particular when Neptune was still far from 30 AU. We cannot exclude this possibility from the point of view of planetary migration. However, we remind the reader that there are several other arguments against collisional grinding in general: (i) the orbital excitation of the cold classical Kuiper belt does not seem to be large enough, compared to that required in the model by Stern and Colwell (1997a); (ii) most of the wide binaries in the cold population would

¹ This is only true if the disk is too excited to support a form of a spiral density wave known as an *apsidal* wave, which can be generated by the v_8 (Ward and Hahn, 1998a, 1998b). Waves such as this would not allow the eccentricity of the individual particles to grow significantly. However, waves will only be generated in disks with e < 0.01, $i < 0.3^{\circ}$ (Hahn, 2003), which is much smaller than that required for collisional grinding to be powerfull enough to deplete 99% of the mass ($e \sim 0.25$, $i \sim 7^{\circ}$, Stern and Colwell, 1997b). So that in a collisional grinding regime the collective response of the disk can be ignored.

not have survived the collisional grinding phase (Petit and Mousis, 2004); (iii) if all conditions for the collisional grinding were met in the Kuiper belt, it is likely that they were met also in the 20–30 AU region, thus preventing the formation of a massive enough Oort cloud of comets (Stern and Weissman, 2001; Charnoz, private communication).

Therefore, we believe that the current location of Neptune and the mass deficiency of the Kuiper belt imply that the proto-planetary disk possessed an edge at about 30 AU. There are at least five mechanisms that could have truncated the disk at small heliocentric distance, prior to planetary accretion: (1) A passing star tidally strips the Kuiper belt after the observed Kuiper belt objects formed (Ida et al., 2000b; Kobayashi and Ida, 2001). (2) An edge formed prior to planetesimal formation due to aerodynamic drag (Youdin and Shu, 2002). (3) An edge formed during planet accretion due to size-dependent radial migration caused by gas drag (Weidenschilling, 2003). (4) Nearby early-type stars photoevaporated the outer regions of the solar nebula before planetesimals could form (Hollenbach and Adams, 2003). (5) Magneto-hydrodynamic instabilities in the outer regions of the disk prevented the formation of planetesimals in these regions (Stone et al., 1998). We stress that the truncation of the disk at ~ 35 AU is *not* in contradiction with the existence of the Kuiper belt beyond 40 AU. In fact, the entire Kuiper belt could have been pushed out from within 35 AU during Neptune's migration, following the mechanisms discussed by Malhotra (1995), Gomes (2003), and Levison and Morbidelli (2003).

6. Migration in a truncated disk

The presence of an edge in a massive disk does not imply that a migrating planet will stop at the edge. Indeed, since angular momentum must be conserved during the migration process, the final location of the planets depends more on the total angular momentum in the disk than on the location of the edge. To illustrate this, Fig. 10 shows Neptune's migration in 6 disks that are initially spread between 10 and 30 AU, but with masses varying from 20 to $100M_{\oplus}$ (all with surface density profile proportional to r^{-1}).² The initial location of Neptune was at 18.1 AU. The disk with $20M_{\oplus}$ has a subcritical surface density. Neptune exhibits a dumped migration and stalls well within the disk. Therefore a massive annulus is preserved between a few AU beyond the planet's location and the original outer edge of the disk. The $30M_{\oplus}$ disk also appears to be initially subcritical and after a fast start, the migration starts to slow. However, a little before $\sim 9 \times 10^7$ yr, there is a brief burst of migration that occurs when Neptune's 3:2 mean motion resonance leaves the disk. We believe that this burst of migration is due to particles



leaving the resonance as explain above. By 2×10^8 yr, there are very few particles left in the disk, although Neptune only reached 27 AU. Interestingly, roughly 2% of the particles can be found in orbits that are decoupled from Neptune beyond 30 AU. Most of these are in mean motion resonances, but some were delivered to this region by the Gomes (2003) mechanism.

The disk with $35M_{\oplus}$ has a surface density close to the critical value. The planet migrates outwards in an almost linear way for ~ 50 Myr. When it reaches ~ 26 AU, the unstable region of the disk (which extends up to a distance of about 1/6th of the planet's semi-major axis; Duncan et al., 1995) reaches the edge of the disk. The planet starts to feel the disk truncation and its migration is rapidly damped. The final location is 2 AU inside the original disk edge, but the entire region beyond the planet has been depleted.

More massive disks have supercritical densities. In the case of $50M_{\oplus}$ the planet stops almost exactly at the disk's edge, while in the other cases it goes several AUs beyond it. We stress that at the end of all our simulations, except the one with $20M_{\oplus}$, the original disk was destroyed despite the fact that the Neptune's final location varied by 8 AU. Therefore, for an observer looking at the final planetary configuration, there would be no way to tell where was the original disk's edge and which was the original mass of the disk. Given a final position of Neptune, there is a one parameter family of solutions for the disk's size and mass that is compatible with the result (assuming a given initial position of the planet; the situation is even more complicated if also the initial position is considered as a free parameter). This is precisely the situation that we are currently facing when we look at our Solar System.

Among the family of possible solutions for the disk's parameters that are compatible with Neptune's location at 30 AU we tend to prefer a mass density close to the critical value of $\sim 1.5 M_{\oplus} \text{ AU}^{-1}$, and an outer edge close to 35 AU.



 $^{^2}$ In order to compare the results of the new integrations with those of Fig. 6, the reader should remind that, for a given total mass, the surface density is 1.6 times higher in the new, narrower disks.

This is due to the fact that the smaller the disk the harder it is to push the observed Kuiper belt objects out to their current locations by the mechanisms of Malhotra (1993, 1995), and particularly by Levison and Morbidelli (2003).

A narrow, low mass disk has several implications concerning events in the inner Solar System. Levison et al. (2001) proposed that a late formation or a late outward migration of Uranus and Neptune triggered the so-called Late Heavy Bombardment of the Moon. To explain the delivery to the Moon of 6×10^{21} g of material, a constraint deduced from models of impact basin formation, they had to postulate that the disk in Uranus–Neptune region contained $100M_{\oplus}$ of planetesimals. As we described above, a disk this massive is inconsistent with the current location of Neptune. However, there are large uncertainties in the total mass of the basin-forming projectiles of at least a factor of a few (see the discussion in Levison et al., 2001), so a disk of $30M_{\oplus}$ might still be compatible with the Late Heavy Bombardment (but it is definitely on the low end).

And finally, the constraints that we have presented in this section on the mass and extent of the original proto-planetary disk has implications for the Solar System formation models presented in Thommes et al. (2003). These authors presented a series of models where the giant planets formed in a very compact configuration that, either during or sometime after Jupiter and Saturn accreted their gaseous envelopes, suffered a dynamical instability that scattered Uranus and Neptune outward. Uranus, Neptune, and perhaps the core of Saturn then had their orbits circularized by the gravitational interaction (i.e., dynamical friction) with the external proto-planetary disk. In their most extreme model, Jupiter and Saturn were fully formed, and Uranus and Neptune were between them, before the instability. We have performed a series of simulations of this extreme case, but where the disk was truncated. In the case were the disk contained $50M_{\oplus}$ between 10 and 35 AU, we found that the probability that both Uranus and Neptune became decoupled from Jupiter and Saturn is smaller than 10% (we did 11 simulations and always lost at least one planet). With a $100M_{\oplus}$ disk we obtained one case out of three simulations where both Uranus and Neptune decoupled from Jupiter and Saturn. However, as the results presented earlier in this section suggest, the outermost planet ended up at 40 AU, too far from the Sun. We caution that in our simulations the disk was represented by only 1000 particles and perhaps the results would be different if the disk was better resolved, although we believe that this is unlikely. Thus, the most extreme version of the Thommes et al. scenario can most likely be ruled out.

7. Migration of planetary embryos

Recall that in Section 5 we performed a study of the Petit et al. (1999) scenario for the mass depletion of the Kuiper belt, where we initially placed an Earth-mass embryo on a circular orbit outside the orbit of Neptune at 27.15 AU.



Fig. 11. The dynamics of a Mars-mass embryo initially placed outside Neptune's orbit (27 and 18 AU, respectively). The disk mass is $30M_{\oplus}$ between 10 and 50 AU. The black curve at the bottom of the panel shows the evolution of Neptune's semi-major axis. The three light grey curves show the evolution of the embryo's perihelion distance, semi-major axis and aphelion distance, respectively from bottom to top. The two curves with intermediate grey color, which evolve parallel to Neptune's semi-major axis, show the location of the 7:9 and 3:4 resonances, respectively. The embryo is initially in the former resonance, and then is capture in the latter at $t \sim 1.5$ Myr. The embryo quits the 3:4 resonance at $t \sim 40$ Myr.

Jupiter, Saturn, Uranus, and Neptune were initially at 5.40, 8.78, 16.25, and 23.14 AU, respectively. In addition, we included a $30M_{\oplus}$ disk between 10 and 50 AU, with a surface density profile decaying as r^{-1} . Naively, we expected that the Earth-mass embryo would have been caught by Neptune during the planet's migration, and subsequently behaved as a scattered disk body (Levison and Duncan, 1997) undergoing repeated close encounters with Neptune, clearing out the Kuiper belt, and eventually being ejected by the giant planets. That, surprisingly, is not what happened.

Figure 8 shows that the embryo migrates much faster than Neptune. In this simulation, the planet migrates very fast to the edge of a disk in a runaway migration that leaves the disk behind, almost un-depleted. Then the embryo reverses the migration, returning to 40 AU, and finally it turns around again, reaching a final position that is well beyond the initial edge of the disk (50 AU). The embryo's final eccentricity and inclinations are $\sim 0.02^{\circ}$ and $\sim 1^{\circ}$, respectively.

Figure 11 shows another example of interesting embryo dynamics. Here the initial conditions were the same as the simulation shown in Fig. 8 except the embryo's mass reduced to that of Mars. The Mars-mass embryo is less mobile than the Earth-mass one, and thus it is more susceptible to be trapped in mean motion resonances. In this case, its eccentricity is first excited to 0.06 by the 7:9 mean motion resonance, which was initially close by. At this value of the eccentricity the 7:9 resonance overlaps with the stronger 3:4 resonance, which captures the embryo at 1.5 Myr. The transition to the new resonance causes the embryo's eccentricity to jump to ~ 0.1 . Once in the resonance two competing effects dominate the embryo's dynamics: the outward mi-

gration of the resonance tends to increase its eccentricity, while the dynamical friction exerted by the disk tends to reduce it. In this case, the dynamical friction slightly dominates, so that the embryo's eccentricity is slowly reduced. At $t \sim 40$ Myr the embryo finally leaves the resonance, and consequently its eccentricity falls dramatically to less than 0.02. The embryo is therefore stabilized outside Neptune's position.

We have performed six simulations like those above varying the number of particles in the disk, the disk mass and the embryo mass. In some runs runaway migration is important while in others it is not. In all cases Neptune stopped before reaching the embryo. Thus, contrary to our expectations, an Earth-mass planetary embryo initially in a low mass disk just outside Neptune's orbit would not have been scattered by Neptune, but would have migrated ahead of Neptune, until—or somewhat beyond—the disk's edge. The embryo would still be present in the Solar System, with a low eccentricity, low inclination orbit which would have not escaped detection in the numerous ecliptic surveys that have been performed for the detection of Kuiper belt bodies.

The reason that the embryos in these simulations could find a stable position well beyond Neptune is that the disk extended to 50 AU. If the disk were truncated at a smaller heliocentric distance, so that Neptune could reach the outer edge, the situation would be drastically different. We performed four simulations to study this situation. In these integrations, we truncated the disk at 30 AU to insure that Neptune will stop near its current location. In half of the cases, Neptune eventually scatters the embryo towards the inner Solar System where it is ejected from the Solar System by the gas giants (Fig. 12). In the remaining cases, Neptune scatters the embryo outwards, where the dynamical friction exerted by the other scattered disk planetesimals damp its eccentricity. The embryo therefore survives on a low-eccentricity orbit, outside the position of Neptune and beyond the original edge of the disk.

In all of the simulations thus far explored, the system consisted of the four giant planets, an embryo, and a disk. We have not studied systems with multiple embryos because this case has already been ruled out by Morbidelli et al. (2001). These authors studied the evolution of a system of multiple embryos initially outside Neptune's orbit and demonstrated that, even neglecting planetary migration and dynamical friction, there are always embryos left on stable orbits beyond Neptune. In these models, the surviving embryos were decoupled from Neptune because of dynamical encounters with other embryos. The inclusion of a transneptunian disk should make the survival of embryos even more likely.

In conclusion, the existence at early epochs of numerous Mars- to Earth-mass embryos outside the primordial position of Neptune seems unlikely. If one such embryo existed, its elimination requires that the primordial massive disk was truncated not far from the current Neptune's position, at 30 AU.



Fig. 12. The evolution of an embryo with $0.2M_{\oplus}$ initially placed outside Neptune (16 and 14 AU, respectively) in a disk with $43M_{\oplus}$ truncated at 30 AU. The black curves show the evolution of the semi-major axes of Jupiter, Saturn, Uranus, and Neptune, from bottom to top, while the light grey curves show the perihelion and aphelion distance of the embryo.

8. Conclusions and discussion

In this paper we have investigated, in detail, the phenomenon of planetary migration due to the scattering of disk planetesimals. Although our explorations cover a much wider parameter space than Hahn and Malhotra (1999), in the region of overlap our results are consistent with theirs.

In the case of the giant planets in our Solar System we have found that—depending on the mass density of the disk—Neptune could have either experienced damped migration (in which case it would have moved only a few AU) or forced migration (which would have driven it to the edge of the disk). However, we also argue that if Neptune experienced damped migration that left a massive Kuiper belt beyond its final position—as proposed by Hahn and Malhotra (1999)—it is difficult to remove this mass, as the current Kuiper belt observations demand, without causing Neptune to migrate too far from the Sun.

Thus, we conclude that the primordial proto-planetary disk was most likely truncated near 30 AU before Neptune arrived on the scene. The exact location of the outer disk edge cannot be determined, because it depends on the disk's mass density. Indeed, in the experiments that we ran with a disk edge at 30 AU and in which the disk was totally depleted, Neptune stopped migrating at distances ranging between 27 and 35 AU. For a number of reasons explained above, our preference is for a disk that extended up to ~ 35 AU, with a linear mass density of about $1.5M_{\oplus}$ AU⁻¹.

We have shown that in very massive disks an isolated Neptune-mass planet can experience a runaway migration that can transport it over very large distances. This process does not require the existence of multiple planets and is selfsustaining, i.e., it occurs because the migration itself feeds particles to the planet that continues to drive the migration. It also can occur in either direction. This phenomenon may be relevant for extra-solar planets. We have also concluded that Earth- or Mars-mass embryos could not have existed in the planetary disk exterior to Neptune unless the disk was truncated. This is additional support for our conclusion that, in order to see the planetary system that we see today, the proto-planetary disk initially must have had an edge at ~ 30 AU.

So far in this paper, we have focused on the evolution of Neptune. Unfortunately, we find that we have a significant problem with Uranus. In all simulations starting from a compact planetary configuration where Neptune is initially inside 20 AU, Uranus always stopped well before its current location at \sim 19 AU. This is because in these cases the planetesimals scattered by Neptune interact with Saturn almost at the same time as they interact with Uranus, so that Uranus effectively 'sees' only a small portion of the total disk's mass. This may indicate that Uranus and Neptune formed at 17-18 and 23-25 AU, respectively (see Hahn and Malhotra, 1999), despite of the apparent difficulty of accreting planets at large heliocentric distances (Levison and Stewart, 2001; Thommes et al., 2003). Alternatively, it may indicate that the migration process was triggered by some instability in the originally compact planetary system, something similar to what was proposed by Thommes et al. (1999). This will be the subject of future investigations.

Acknowledgments

R.G. thanks CNPq and FAPERJ for supporting grants. He is also indebted to R. Vieira Martins for allowing the use his personal computers for the numerical integrations. A.M. and H.F.L. are grateful to CNRS and NSF for supporting H.F.L.'s sabbatical at the Nice Observatory. H.F.L also thanks NASA's *Origins* program for supporting his involvement in this research.

Appendix A. Integration methods

The simulations presented in this paper have been performed with two different numerical integrators. In both integration schemes, the disk planetesimals interact with the planets but not with each-other, which significantly reduces the simulation time.

The simulations illustrated in Sections 3 and 4 have used the MERCURY integrator (Chambers, 1999), with a timestep of one year and a relative accuracy during encounters equal to 10^{-10} . Disk particles were discarded when they reached a heliocentric distance of 1000 AU. For the simulations in Section 4 a 'trick' was used that decreased CPU time, while retaining the resolution of a 10,000 particle disk. This trick is based on the idea that, at any time in our simulations, a 10,000 particle resolution is required only for the fraction of the disk that is suffering encounters with Neptune. A lower resolution is adequate for the region of the disk that has yet to get close to Neptune. Thus, our disks are initially made of only 1000 objects. A particle is replaced by ten objects (clones) with one tenth of the original body's mass and slightly different coordinates, when it first evolves onto an orbit with a semi-major axis less than $a_N + 2R_H$ (where a_N and R_H are the Neptune's semi-major axis and the Hill's radius). Because, in general, the dynamics of bodies in this type of orbit are chaotic, the orbital evolution of the clones rapidly diverged. Thus, the clones experience totally independent close encounters with Neptune. With this trick we could simulate Neptune's migration with a resolution of a 10,000 particles, by integrating at most 2500 planetesimals at any time.

The simulations reported in Sections 5 and 6 have been done using the integrator SyMBA (Duncan et al., 1998). We used 4700 particles to simulate the disk in the simulations of Neptune's migration in presence of an embryo presented in Section 5. The same has been done for the simulation of the dynamics of the martian-mass embryo in Section 7. The simulations in Section 6 on Neptune's migration in presence of a disk's edge have been done with 10,000 disk particles. SyMBA integrator has also been used to reproduce the simulations of runaway migration discussed in Section 3, leading to the reversal of Neptune's motion.

In the SyMBA integrations we discarded disk particles when they became closer than 4.5 AU to the Sun. This has been done to speed up the simulation, use a larger time-step, and avoid the problem of the accuracy of the integration of particles with small perihelion distance (Levison and Duncan, 2000). The elimination of a significant number of particles as they enter the inner Solar System significantly affects the migration of Jupiter, which consequently, proceeds outwards. This artifact, however, presumably has no impact on the migration of Neptune, on which this work is focused.

References

- Binney, J., Tremaine, S., 1987. Galactic Dynamics. Princeton Univ. Press, Princeton, NJ. 747 p.
- Brown, M., 2001. The inclination distribution of the Kuiper belt. Astron. J. 121, 2804–2814.
- Chambers, J.E., 1999. A hybrid symplectic integrator that permits close encounters between massive bodies. Mon. Not. R. Astron. Soc. 304, 793–799.
- Chiang, E.I., Brown, M.E., 1999. Keck pencil-beam survey for faint Kuiper belt objects. Astron. J. 118, 1411–1422.
- Davis, D.R., Farinella, P., 1997. Collisional evolution of Edgeworth–Kuiper belt objects. Icarus 125, 50–60.
- Davis, D.R., Farinella, P., 1998. Collisional erosion of a massive Edgeworth–Kuiper belt: constraints on the initial population. In: Proc. Lunar Planet. Sci. Conf. 29th, pp. 1437–1438.
- Duncan, M.J., Levison, H.F., Budd, S.M., 1995. The long-term stability of orbits in the Kuiper belt. Astron. J. 110, 3073–3083.
- Duncan, M.J., Levison, H.F., Lee, M.H., 1998. A multiple time step symplectic algorithm for integrating close encounters. Astron. J. 116, 2067– 2077.
- Fernandez, J.A., Ip, W.H., 1984. Some dynamical aspects of the accretion of Uranus and Neptune: the exchange of orbital angular momentum with planetesimals. Icarus 58, 109–120.

- Gladman, B., Kavelaars, J.J., Petit, J.M., Morbidelli, A., Holman, M.J., Loredo, Y., 2001. The structure of the Kuiper belt: size distribution and radial extent. Astron. J. 122, 1051–1066.
- Goldreich, P., Tremaine, S., 1980. Disk-satellite interactions. Astrophys. J. 241, 425–441.
- Gomes, R.S., 2003. The origin of the Kuiper belt high inclination population. Icarus 161, 404–418.
- Grogan, K., Dermott, S.F., Durda, D.D., 2001. The size-frequency distribution of the zodiacal cloud: evidence from the Solar System dust bands. Icarus 152, 251–267.
- Gustafson, B.A.S., 1994. Physics of zodiacal cloud. Annu. Rev. Earth Planet. Sci. 22, 553–595.
- Hahn, J.M., 2003. The secular evolution of the Kuiper belt. Astrophys. J. 595, 531–549.
- Hahn, J.M., Malhotra, R., 1999. Orbital evolution of planets embedded in a planetesimal disk. Astron. J. 117, 3041–3053.
- Hayashi, C., 1981. Structure of the solar nebula, growth and decay of magnetic fields and effects of magnetic and turbulent viscosities on the nebula. Prog. Theor. Phys. Suppl. 70, 35–53.
- Hollenbach, D., Adams, F.C., 2003. Dispersal of disks around young stars: constraints on Kuiper belt formation. In: Caroff, L., Backman, D. (Eds.), Debris Disks and the Formation of Planets. ASP, San Francisco.
- Holman, M.J., Wisdom, J., 1993. Dynamical stability in the outer Solar System and the delivery of short period comets. Astron. J. 105, 1987– 1999.
- Ida, S., Larwood, J., Burkert, A., 2000b. Evidence for early stellar encounters in the orbital distribution of Edgeworth–Kuiper belt objects. Astrophys. J. 528, 351–356.
- Ida, S., Bryden, G., Lin, D.N.C., Tanaka, H., 2000b. Orbital migration of Neptune and orbital distribution of trans-neptunian objects. Astrophys. J. 534, 428–445.
- Jewitt, D., Luu, J., Chen, J., 1996. The Mauna–Kea–Cerro–Totlolo (MKCT) Kuiper belt and Centaur survey. Astron. J. 112, 1225–1232.
- Kenyon, S.J., Luu, J.X., 1999. Accretion in the early outer Solar System. Astrophys. J. 526, 465–470.
- Kobayashi, H., Ida, S., 2001. The effects of a stellar encounter on a planetesimal disk. Icarus 153, 416–429.
- Levison, H.F., Duncan, M.J., 1997. From the Kuiper belt to Jupiter-family comets: the spatial distribution of ecliptic comets. Icarus 127, 13–32.
- Levison, H.F., Duncan, M.J., 2000. Symplectically integrating close encounters with the Sun. Astron. J. 120, 2117–2123.
- Levison, H.F., Morbidelli, A., 2003. The formation of the Kuiper belt by the outward transport of bodies during Neptune's migration. Nature 426, 419–421.
- Levison, H.F., Stewart, G.R., 2001. Remarks on modeling the formation of Uranus and Neptune. Icarus 153, 224–228.
- Levison, H.F., Dones, L., Chapman, C.R., Stern, S.A., Duncan, M.J., Zhanle, K., 2001. Could the lunar late heavy bombardment have been triggered by the formation of Uranus and Neptune? Icarus 153, 111– 129.
- Liou, J., Zook, H.A., 1999. Signatures of the giant planets imprinted on the Edgeworth–Kuiper belt dust disk. Astron. J. 118, 580–590.
- Liou, J., Zook, H.A., Dermott, S.F., 1996. Kuiper belt dust grains as a source of interplanetary dust particles. Icarus 124, 429–440.
- Malhotra, R., 1993. The origin of Pluto's peculiar orbit. Nature 365, 819–821.
- Malhotra, R., 1995. The origin of Pluto's orbit: implications for the Solar System beyond Neptune. Astron. J. 110, 420–429.
- Masset, F.S., 2001. On the co-orbital corotation torque in a viscous disk and its impact on planetary migration. Astrophys. J. 558, 453–462.

- Morbidelli, A., Brown, M., 2004. The Kuiper belt and the primordial evolution of the Solar System. In: Festou, M., Keller, H.U., Weaver, H. (Eds.), Comet II. Univ. of Arizona Press, Tucson, AZ. In press.
- Morbidelli, A., Valsecchi, G.B., 1997. Neptune scattered planetesimals could have sculpted the primordial Edgeworth–Kuiper belt. Icarus 128, 464–468.
- Morbidelli, A., Jacob, C., Petit, J.M., 2001. Planetary embryos never formed in the Kuiper belt. Icarus 157, 241–248.
- Moro-Martín, A., Malhotra, R., 2003. Dynamical models of Kuiper belt dust in the inner and outer Solar System. Astron. J. 125, 2255–2265.
- Murray, N., Hansen, B., Holman, M., Tremaine, S., 1998. Migrating planets. Science 279, 69.
- Nagasawa, M., Ida, S., 2000. Sweeping secular resonances in the Kuiper belt caused by depletion of the solar nebula. Astron. J. 120, 3311–3322.
- Öpik, E.J., 1976. Interplanetary Encounters. Elsevier, New York.
- Ozernoy, L.M., Gorkavyi, N.N., Mather, J.C., Taidakova, T.A., 2000. Signatures of exosolar planets in dust debris disks. Astrophys. J. 537, L146– L151.
- Petit, J.M., Mousis, O., 2004. KBO binaries: how numerous were they? Icarus 168, 409–419.
- Petit, J.M., Morbidelli, A., Valsecchi, G.B., 1999. Large scattered planetesimals and the excitation of the small body belts. Icarus 141, 367– 387.
- Stern, S.A., Colwell, J.E., 1997a. Accretion in the Edgeworth–Kuiper belt: forming 100–1000 km radius bodies at 30 AU and beyond. Astron. J. 114, 841–849.
- Stern, S.A., Colwell, J.E., 1997b. Collisional erosion in the primordial Edgeworth–Kuiper belt and the generation of the 30–50 AU Kuiper gap. Astrophys. J. 490, 879–885.
- Stern, S.A., Weissman, P.R., 2001. Rapid collisional evolution of comets during the formation of the Oort cloud. Nature 409, 589–591.
- Stone, J.M., Gammie, C.F., Balbus, S.A., Hawley, J.F., 1998. Transport processes in protostellar disks. In: Mannings, V., Boss, A.P., Russell, S.S. (Eds.), Protostars and Planets IV. Univ. of Arizona Press, Tucson, AZ, pp. 589–612.
- Thommes, E.W., Duncan, M.J., Levison, H.F., 1999. The formation of Uranus and Neptune in the Jupiter–Saturn region of the Solar System. Nature 402, 635–638.
- Thommes, E.W., Duncan, M.J., Levison, H.F., 2003. Oligarchic growth of giant planets. Icarus 161, 431–455.
- Trujillo, C.A., Jewitt, D.C., Luu, J.X., 2001. Properties of the trans-neptunian belt: statistics from the Canada–France–Hawaii telescope survey. Astron. J. 122, 457–473.
- Valsecchi, A., Manara, G.B., 1997. Dynamics of comets in the outer planetary region. II. Enhanced planetary masses and orbital evolutionary paths. Astron. Astrophys. 323, 986–998.
- Wahhaj, Z., Koerner, D.W., Ressler, M.E., Werner, M.W., Backman, D.E., Sargent, A.I., 2003. The inner rings of β Pictoris. Astrophys. J. 584, L27–L31.
- Ward, W.R., 1997. Protoplanet migration by nebula tides. Icarus 126, 261– 281.
- Ward, W.R., Hahn, J.M., 1998a. Neptune's eccentricity and the nature of the Kuiper belt. Science 280, 2104–2107.
- Ward, W.R., Hahn, J.M., 1998b. Dynamics of the trans-Neptune region: apsidal waves in the Kuiper belt. Astron. J. 116, 489–498.
- Weidenschilling, S., 2003. Formation of planetesimals/cometesimals in the solar nebula. In: Festou, M., Keller, H.U., Weaver, H. (Eds.), Comet II. Univ. of Arizona Press, Tucson, AZ.
- Youdin, A.N., Shu, F.H., 2002. Planetesimal formation by gravitational instability. Astrophys. J. 580, 494–505.