Dynamics of Populations of Planetary Systems
Proceedings IAU Colloquium No. 197, 2005
Z. Knežević and A. Milani, eds.

DOI: 10.1017/S1743921304008798

Interaction of planetesimals with the giant planets and the shaping of the trans-Neptunian belt

Harold F. Levison[1] **and Alessandro Morbidelli**[2]

[1]Southwest Research Institute, 1050 Walnut St., Suite 400, Boulder, CO 80302, USA
email: hal@boulder.swri.edu

[2]Obs. de la Côte d'Azur, B.P. 4229, 06034 Nice Cedex 4, France

Abstract. The trans-Neptunian region is inhabited by multiple dynamical populations, each of which have a complicated structure. For the most part, these structures cannot have been sculpted by the giant planets, once on their current orbital configuration. Thus, they represent important clues to the conditions that existed in the distant past. We argue in this paper, that most of what we see is the result of the outward migration of Neptune. By combining results from various authors, we can reproduce most of the observed properties of the trans-Neptunian region. Several aspects are not yet totally clear, and some may not be totally correct. But, for the first time, we have a view — if not not a detailed model — of how the system formed.

Keywords. Kuiper Belt, solar system: formation, solar system: general, comets: general

1. Introduction

It has been over a decade since the first object beyond the orbit of Neptune was discovered (Jewitt & Luu 1993). Since that time over 850 so called trans-Neptunian objects (or TNOs) have been discovered, and over 400 of these have been observed often enough so that their orbits are reliable†. As these data were collected, it became clear that rather than being the dynamically cold, smooth, and basically uninteresting system that most researchers expected, the trans-Neptunian region is home to a complex series of overlapping dynamical structures, each with a complicated, and sometimes diffcult to understand, configuration.

This complexity has been a significant boon to the planetary science community. In addition to keeping us funded, the trans-Neptunian region represents an excellent site to study the early dynamical evolution of the planetary system. This is due to the fact that, for the most part, the TNOs that we see cannot have formed in the orbits in which we see them in. Indeed, they must have accreted on low-inclination, nearly circular orbits (Stern 1996). In addition, they cannot have been placed on their current orbits by the planets in their current configuration. However, something must have moved the TNOs from their primordial orbits to the ones we observe, and thus, the outer planetary system must have looked very different in the distant past.

The current TNO orbits are, therefore, a diagnostic to what happened to the outer Solar System during the epoch when the planets formed. And just like a forensic scientists at a crime seen can determine the perpetrator of the crime, we believe that the dynamical structure of the trans-Neptunian region will reveal to us the mysteries of the birth of the Solar System. In this paper, we will discuss the current status of our investigation. In §2

---

† For more information see `http://cfa-www.harvard.edu/cfa/ps/lists/TNOs.html` and `http://cfa-www.harvard.edu/iau/lists/Centaurs.html`.

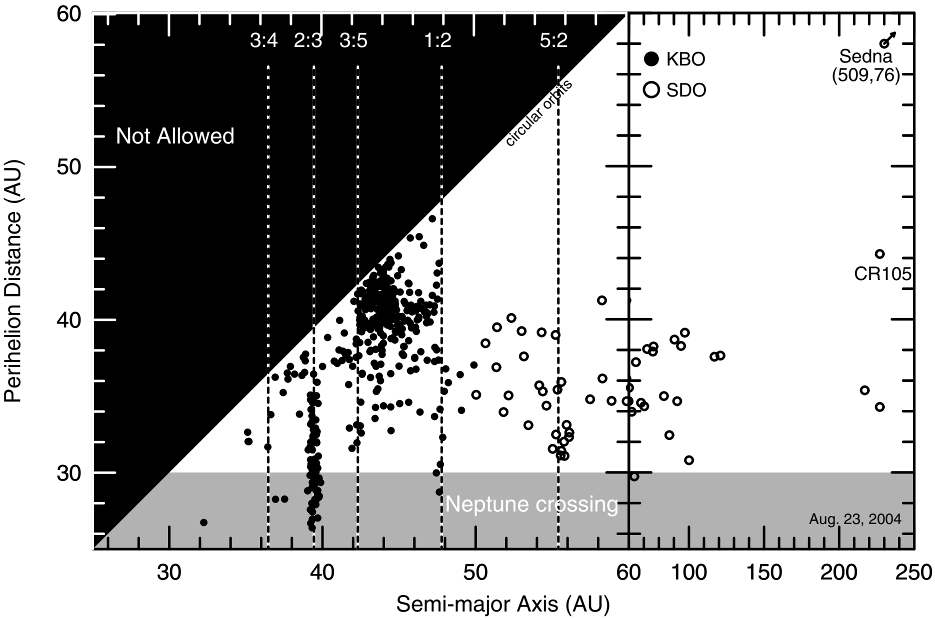

Figure 1. The semi-major axis – perihelion distance distribution for all the known TNOs with multi-opposition observations as of August 23, 2004. The filled and open dots represent the KBOs and SDOs, respectively, as determined my the Minor Planet Center. The dashed lines show a few of the mean motion resonances with Neptune. The black region is forbidden because perihelion distance cannot be larger than the semi-major axis. The gray region marked orbits that cross the orbit of Neptune.

we discuss the observed dynamical features and sub-populations of the trans-neptunian region. In §3 we briefly review planet migration. We discuss the models for the origin of these sub-populations in §4 – §9. We conclude in §10.

2. The Observed Structure of the Trans-Neptunian Region

Figure 1 shows the semi-major axis (a) – perihelion distance (q) distribution for all the TNOs that have been observed over multiple oppositions as of August 23, 2004. The figure shows the orbits of these objects as dots. The location of Neptune's important mean motion resonances are shown by vertical dashed lines. Objects in the gray region cross the orbit of Neptune. In general, they are stabilized because they are trapped in a mean motion resonance with Neptune.

The trans-Neptunian population is typically divided into two distinct structures, called the *Kuiper belt* and the *scattered disk*. Traditionally (i.e. four or five years ago), the scattered disk was thought to consist of objects that were placed on their current orbits by the giant planets in their current configuration (Duncan & Levison 1997). That is, they were thought to be objects that were scattered by Neptune onto orbits with large semi-major axes and large eccentricities. As a result, they should be on orbits that pass close to Neptune's orbit (i.e. with perihelion distance q less than ~ 35 AU). If this idea were true, they would represent a dynamically active population where objects are slowly

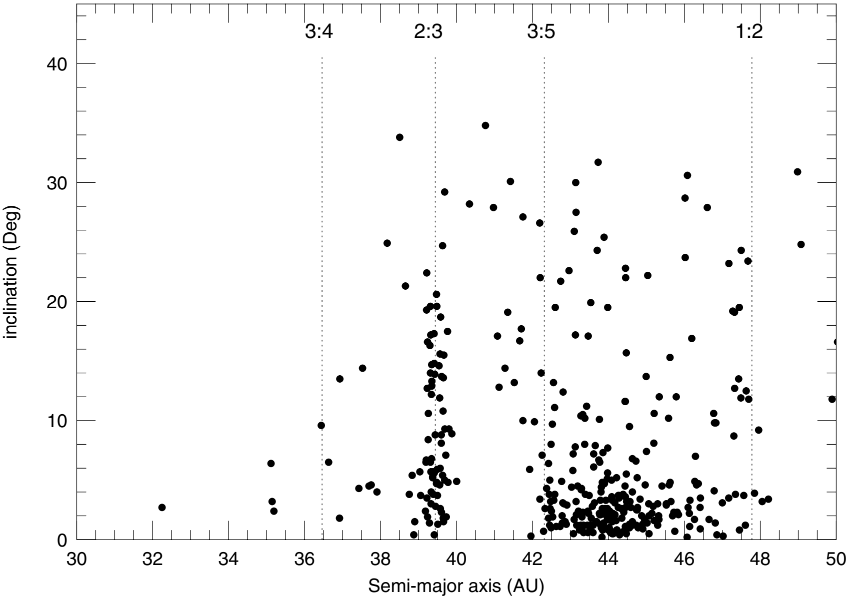

Figure 2. The semi-major axis – inclination distribution for all the known $a < 50$ AU TNOs with multi-opposition observations as of August 23, 2004. The dashed lines show a few of the mean motion resonances with Neptune.

diffusing through orbital element space, and thus the bodies that belong in it would not provide us many direct clues to uncover the primordial architecture of the Solar System. As we describe more in details below, this definition is not strictly true, and so the scattered disk is the source of some important constraints on planet formation.

The *Kuiper belt* is basically the complement of the scattered disk in the $a > 30$ AU region. As a result, it consists of objects which are essentially "frozen" in their current orbits (Duncan *et al.* 1995). As described above, all bodies in the Solar System must have been formed on orbits typical of an accretion disk (e.g. with very small eccentricities and inclinations; Stern 1996). Therefore, all the properties that make the current Kuiper belt very different from an accretion disk must have been acquired during a primordial phase of the Solar System.

In Figure 1, scattered disk objects (SDOs) and the Kuiper belt objects (KBOs) are represented by open and filled circles, respectively. However, to generate this figure we employed the classification scheme used by the Minor Planet Center, which simply calls any object with semi-major axis a smaller than 50 AU a KBO, while objects with $a >$ 50 AU are classified as SDOs. This scheme, of course, has the advantage that it does not require a detailed understanding of the behavior of TNOs. However, it has very little physical meaning. Having said this, the distinction between classes, as shown in the figure, is probably appropriate. The only major change that we would recommend is to consider the non-resonant objects with $40 \lesssim a \lesssim 50$ AU and $q \lesssim 39$ AU SDOs rather than KBOs.

Any theory of the early evolution of the outer Solar System must explain the following characteristics of the trans-Neptunian region:

(*a*) *The existence of resonant populations.* A glance at Figure 1 shows that the Kuiper belt is made of two distinct sub-populations: the *resonant population* and the *classical belt.* The former is made of the objects located in some major mean motion resonance with Neptune (essentially the 3:4, 2:3 and 1:2 resonances, but also the 2:5), while the classical belt objects are not in any noticeable resonant configuration. According to Trujillo *et al.* (2001) the resonant population accounts for only 10% of the total Kuiper belt population, when observational biases are corrected for.

(*b*) *The bi-modal inclination distribution of the classical belt.* Figure 2 shows that the bodies in the Kuiper belt can have very large inclinations. An analysis of observational biases by Brown (2001) suggests that the real inclination distribution of the non-resonant classical belt is bimodal (this is particularly noticeable in the region between the 2:3 and 1:2 mean motion resonance). About half of the objects have an inclination smaller than 4°, while the remaining half have a very broad inclination distribution ranging up to 30° or more. We will refer to these two groups as the *cold* and the *hot* populations, respectively. Interestingly, these two populations seem to have different physical properties. The largest Kuiper belt objects are all in the hot population (Levison & Stern 2001; Trujillo & Brown 2003); moreover there is a statistically significant difference between the color distributions within the two populations (Tegler & Romanishin 2000; Doressoundiram *et al.* 2001; Trujillo & Brown 2002).

(*c*) *The outer edge of the classical belt.* It was generally expected that the number of the Kuiper belt objects should smoothly change with heliocentric distance. However the persisting lack of detection of objects beyond about 50 AU (Allen *et al.* 2001, 2002) cannot be explained by observational biases, but implies a statistically significant steep drop off in number density of large objects (Trujillo & Brown 2001). Figure 1 suggests that the edge of the classical belt coincides with the location of the 1:2 mean motion resonance with Neptune. In addition, there appears to be a correlation between eccentricity and semi-major axis in the classical Kuiper belt. There are many objects with $e \sim 0$ interior to 45 AU, but beyond $a = 45$ AU e on average increases with a.

(*d*) *Existence of an 'extended scattered disk' population.* Gladman *et al.* (2002) was the first to point out that there exists a population of SDOs (i.e. objects with $a > 50$ AU and highly eccentric orbits) that have large enough perihelion distances that they are decoupled from Neptune. The first obvious example of this population was 2000 CR_{105} ($a = 230$ AU and $q = 44.17$ AU), however, Emel'yanenko *et al.* (2003) have shown that several others exist. The most extreme is (90377) Sedna, with a semi-major axis of 509 AU and a perihelion distance of 76 AU (Brown et al. 2004). We call these objects *extended scattered disk* objects because they are on orbits with semi-major axes similar to other scattered disk objects, but their perihelion distances are outside (or 'extended' beyond) the range for the normal scattered disk (Duncan & Levison, 1997, Gladman *et al.* 2001; Emel'yanenko *et al.* 2003). Their large eccentricities strongly suggest that they were gravitationally scattered onto their current orbits. However, this cannot have been done by the current planetary system.

(*e*) *The mass of the Kuiper belt.* The current mass of the Kuiper belt in the 40–50 AU region is estimated from detection statistics to be of order 0.1 Earth masses ($M_\oplus$) only (Jewitt *et al.* 1996, Chiang & Brown 1999, Trujillo *et al.* 2001, Gladman *et al.* 2001). However, models of the accretion process argue that there should have been several tens of Earth masses in this region, in order to produce the observed objects within a reasonable time interval of several 10^7 to 10^8 years (Stern, 1996; Stern & Colwell, 1997a; Kenyon and Luu, 1998, 1999a, 1999b). What happened to this mass is a critical issue. We refer to this as the *mass deficit problem*.

A large number of mechanisms have been proposed so far to explain some of the above mentioned properties of the Kuiper belt. For space limitation we debate here only those which in our opinion –at the light of our current observational knowledge of the Kuiper belt– played a role in the primordial sculpting of the trans-Neptunian population. A more exhaustive review can be found in Morbidelli and Brown (2003).

3. Giant planet migration

It is now well accepted that, after their formation, Uranus and Neptune underwent substantial outward migration due to the scattering of the planetesimals formed in their vicinity (Hahn & Malhotra 1999; Gomes *et al.* 2004). Planet formation is not expected to be 100% efficient. As the masses of Uranus and Neptune approached their full adult values, they must have gravitationally scattered most of the remaining nearby planetesimals both outward into the trans-Neptunian region and inward toward Jupiter and Saturn. Usually the objects scattered outward returned to the Uranus-Neptune region in the next orbit to be scattered again (either inward or outward). Jupiter and Saturn are so massive, however, that they can efficiently eject many of these planetesimals out of the planetary system.

Thus, the objects scattered outward usually return to Uranus and Neptune, while those scattered inward are effectively removed, thereby causing a net flux of objects toward Jupiter. This inward transport of mass required an outward migration of the orbits of Uranus and Neptune in order to conserve the angular momentum and energy of the system. For the same reason, Jupiter migrated inward.

We will argue in the following sections that much of the observed structure in the trans-Neptunian region is the direct result of the outward migration of Neptune. This is particularly true for the Kuiper belt. Indeed, we suggest that all of the observed KBOs were pushed onto their current orbits during this process.

4. Origin of the resonant populations

As we described in the last section, Neptune should have migrated outwards during the final stages of planet formation. Malhotra (1993, 1995) realized that, following Neptune's migration, the mean motion resonances with Neptune also migrated outwards, sweeping the primordial Kuiper belt until they reached their present position. From adiabatic theory (Henrard 1982), some of the Kuiper belt objects swept by a mean motion resonance would have been captured into resonance; they would have subsequently been push outward with the resonance in its migration, which would in turn has driven up their eccentricities. This model accounts for the existence of the large number of Kuiper belt objects in the 2:3 mean motion resonance with Neptune (and also in other resonances) and explains their large eccentricities (see Figure 3).

The mechanism of adiabatic capture into resonance requires that Neptune's migration happened very smoothly. If Neptune had encountered a significant number of large bodies (Lunar mass or more), its jerky migration would have jeopardized capture into resonances. Hahn & Malhotra (1999), who simulated Neptune's migration using a disk of Lunar to Martian-mass planetesimals, did not obtain any permanent capture.

Reproducing the observed range of eccentricities of the resonant bodies requires that Neptune migrated by at least 7 AU. Malhotra's simulations also showed that the bodies captured in the 2:3 resonance can acquire large inclinations, comparable to that of Pluto and other objects. The mechanisms that excite the inclination during the capture

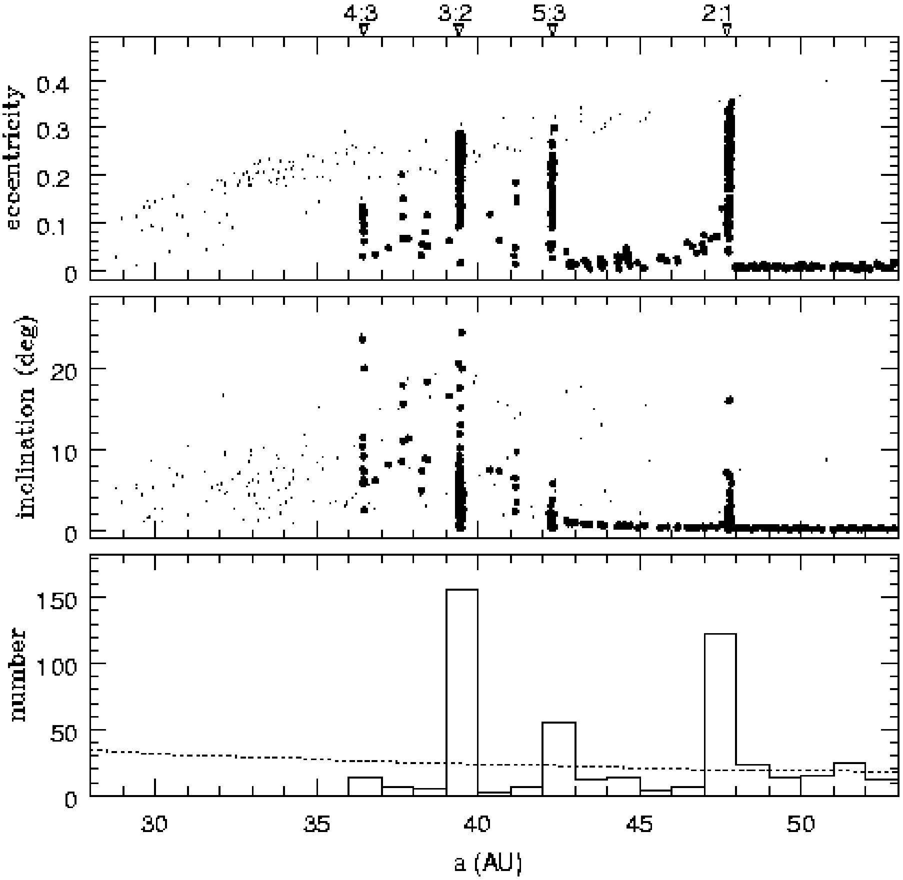

Figure 3. Top and middle panel: Final distribution of the Kuiper belt bodies according to the sweeping resonances scenario (courtesy of R. Malhotra). The simulation is done by numerical integrating, over a 200 Myr timespan, the evolution of 800 test particles on initial quasi–circular and coplanar orbits. The planets are forced to migrate (Jupiter: -0.2 AU;, Saturn: 0.8 AU; Uranus: 3 AU; Neptune: 7 AU) and reach their current orbits on an exponential timescale of 4 Myr. Large solid dots represent 'surviving' particles (i.e., those that have not suffered any planetary close encounters during the integration time); small dots represent the 'removed' particles at the time of their close encounter with a planet. Bottom panel: Histogram of semi-major axis distribution of the surviving bodies. The dashed curve denotes the initial distribution.

process have been investigated in detail by Gomes (2000), who concluded that, although large inclinations can be achieved, the resulting proportion between the number of high inclination vs. low inclination bodies and their distribution in the eccentricity vs. inclination plane do not reproduce well the observations. According to Gomes (2003) most high inclination Plutinos were captured during Neptune's migration from the scattered disk population, rather than from an originally cold Kuiper belt as in Malhotra scenario. We discuss Gomes's scenario in more detail in the next section.

5. Origin of the hot population

Gomes (2003) showed that the Neptune migration, in addition to the resonant populations, can also explain the origin of the hot classical Kuiper belt. Like Hahn & Malhotra

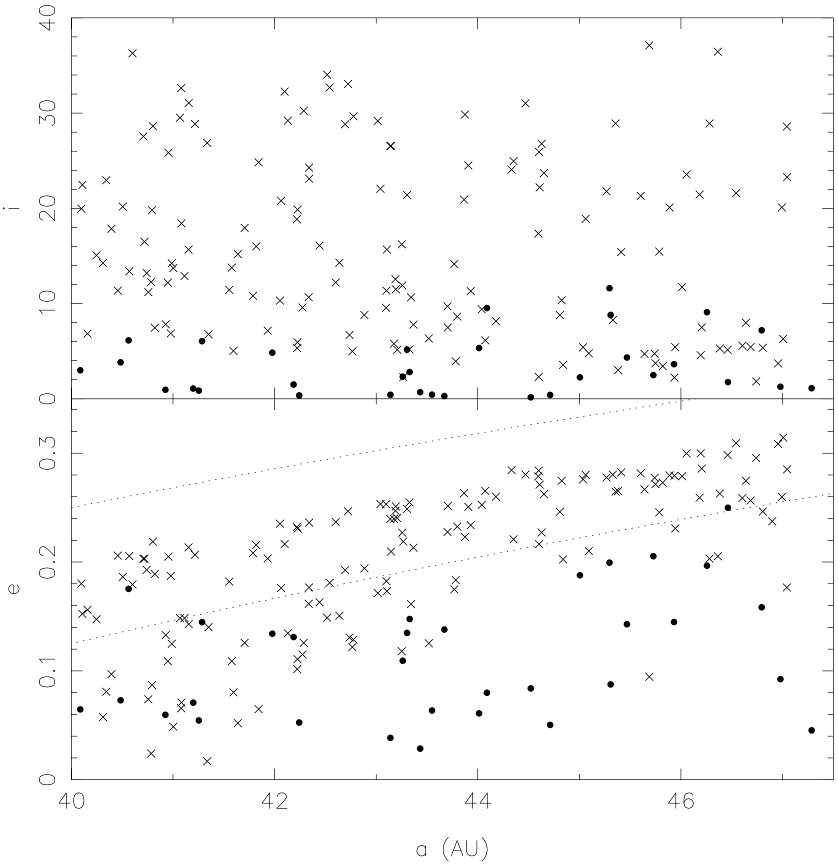

Figure 4. The orbital distribution in the classical belt according to Gomes' simulations. The dots denote a local population, which is only moderately dynamically excited as the mean motion resonances swept by. The crosses denote the bodies that were originally inside 30 AU and were trapped in the Kuiper belt via the mechanism described in the text. Note that the scattered population has a high inclination distribution, that is similar to the observed dynamically hot population. The dotted curves in the eccentricity vs. semi-major axis plot correspond to $q = 30$ AU and $q = 35$ AU. Courtesy of R. Gomes.

(1999), Gomes simulated Neptune's migration, starting from about 15 AU, by the interaction with a massive planetesimal disk with an inner edge slightly beyond Neptune's initial position. However, taking advantage of improved computer technology, he used 10,000 particles to simulate the disk population, with individual masses roughly equal to twice the Pluto's mass. This is in contrast to Hahn & Malhotra, who used only 1,000 particles, with Lunar to Martian masses.

He found that during its migration Neptune scattered planetesimals outward, forming a massive scattered disk. Some of the scattered bodies decoupled from the planet, by decreasing their eccentricity through the interaction with some secular or mean-motion resonance. If Neptune had not been migrating, the decoupled phases would have been transient, because the objects would have remained in the resonance and thus the process is reversible. However, Neptune's migration broke the reversibility because the resonances moved. Thus, some of the decoupled bodies managed to escape from the resonances, and remained permanently trapped in the Kuiper belt.

Since the particles that end up on stable Kuiper belt orbits had scattered off Neptune during their dynamical evolution, they tend to have an extended inclination distribution that is consistent with the observed hot population (see Figure 4). Recall that the inclination distribution of the Kuiper belt is bimodal. Gomes argued that the observations are well matched if it is assumed that the cold population is primordial and his mechanism produced the hot population. While this idea may be consistent with the observed inclination distribution, there are other reasons to believe that the cold population is not primordial, which we discuss in §7. However, we do think that Gomes's mechanism is the most likely cause for the dynamically hot population of the Kuiper belt.

Assuming that the bodies' color varied in the primordial disk with heliocentric distance, Gomes scenario also explains why the scattered objects, and hot classical belt objects, which mostly come from regions inside ∼30 AU, appear to have similar color distributions, while the cold classical objects –the only ones that actually formed in the trans-Neptunian region– have a different color distribution. The Plutinos would be a mixture of the two populations. Similarly, assuming that the maximal size of the objects was a decreasing function of the heliocentric distance at which they formed, Gomes scenario also explains why the biggest Kuiper belt objects are all at large inclination.

Gomes's scenario also helps solve the mass depletion problem. In his simulations only ∼0.2% of the bodies initially in the disk swept by Neptune remained in the Kuiper belt on stable high-i orbits at the end of Neptune's migration. This naturally explains the current low mass of the hot population. However, if the dynamically cold Kuiper belt is primordial, as Gomes believes, then the mass depletion problem still remains for this population.

6. Origin of the outer edge of the Kuiper belt

The existence of an outer edge to the Kuiper belt is, perhaps, its most intriguing characteristic. At least five mechanisms for its origin have been proposed, none of which has yet the general consensus of the community of the experts:

(*a*) A passing star tidally strips the Kuiper belt after the observed Kuiper belt objects formed (Ida *et al.* 2000; Kobayashi & Ida 2001; Melita *et al.* 2002).

(*b*) An edge formed prior to planetesimal formation due to aerodynamic drag (Youdin & Shu 2002).

(*c*) An edge formed during planet accretion due to size-dependent radial migration caused by gas drag (Weidenschilling 2003).

(*d*) Nearby early-type stars photo-evaporated the outer regions of the solar nebula before planetesimals could form (Hollenbach & Adams 2003).

(*e*) Magneto-hydrodynamic instabilities in the outer regions of the disk prevented the formation of planetesimals in these regions (Stone *et al.* 1998).

Unfortunately, space limitations do not allow us to discuss each of these in detail. However, a couple of points need to be made. Firstly, many of these mechanisms truncate the disk of planetesimals, from which the KBOs/SDOs form, without affecting the size of the solar nebula. Thus, observations of the sizes of extra-solar disks are not necessarily a direct constraint on where the edge of the Kuiper belt should lie.

In addition, more importantly to the current discussion, none of these mechanisms predict where the edge of the Kuiper belt should lie. In particular, none can explain why the outer edge of the Kuiper belt lies at the location of the 1:2 resonance. Either we must conclude that this is a coincidence, or we are still missing an important piece of the puzzle. We return to this issue again in the next section.

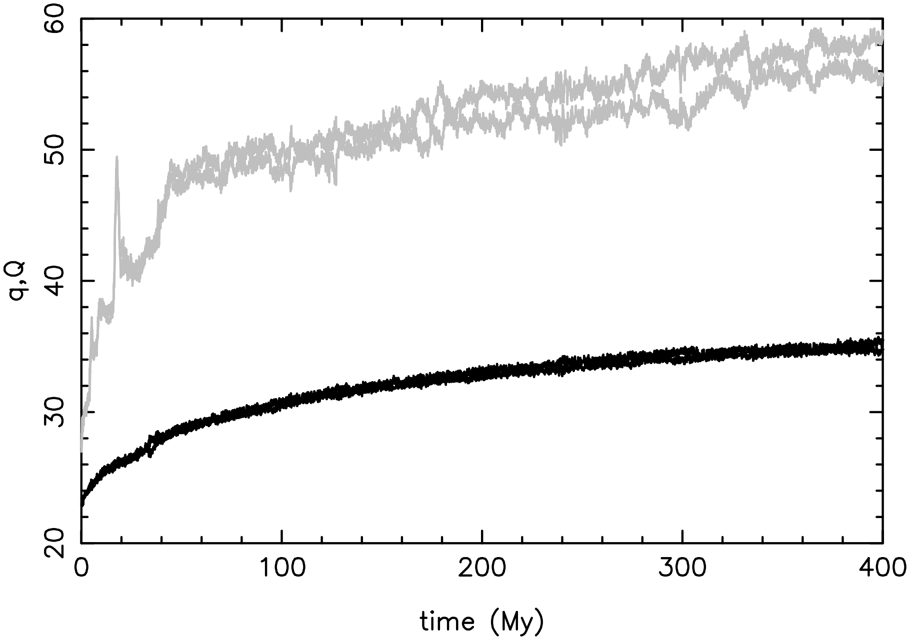

Figure 5. A self-consistent simulation of the Petit *et al.* (1999) scenario for the excitation and dynamical depletion of the Kuiper belt (from Gomes *et al.* 2004). Neptune is originally assumed at ~ 23 AU and an Earth-mass embryo at ~ 27 AU. Both planets are embedded in a 30 $M_\oplus$ disk, extending from 10 to 50 AU with a r^{-1} surface density profile (7.5 $M_\oplus$ between 40 and 50 AU). The black curve shows the evolution of Neptune's heliocentric distance, while the pair of gray curves refer to the perihelion and aphelion distances of the embryo. Notice that the embryo is never scattered by Neptune, unlike in Petit *et al.* simulations. It migrates through the disk faster than Neptune until the disk's outer edge. Neptune interacts with the entire 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 (it reaches 40 AU after 1 Gy).

7. The mass depletion of the cold population

As described in §5, the mechanism proposed by Gomes (2003) for the origin of the dynamically hot classical Kuiper belt naturally explains the mass depletion for this population. However, the population originally in the 40-50 AU region — which eventually became the cold population in Gomes scenario — would have been only moderately excited and not dynamically depleted during Neptune's migration, so that it should have preserved most of its primordial mass. This is still a significant problem in our understanding of the Kuiper belt, because of the observed low mass in this region.

Two general mechanisms have been proposed for the mass depletion: the dynamical ejection of most of the bodies from the Kuiper belt to the Neptune-crossing orbits and the collisional grinding of most of the mass of the Kuiper belt into dust.

All dynamical depletion mechanisms suffer from the same problem that is perhaps best illustrated by an example. The first such mechanism was proposed by Morbidelli & 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 a high-eccentricity orbit that crossed the Kuiper belt for ~ 10^8 y. The repeated passage of the embryo through the Kuiper belt excited the eccentricities of the Kuiper belt bodies, the vast majority of which became Neptune crosser and were subsequently dynamically eliminated by the planets' scattering action. In the Petit *et al.* (1999) integrations that supported this scenario, however, the Kuiper belt bodies were treated as test particles, and therefore their ejection to Neptune-crossing orbit did not alter the position of Neptune.

Gomes *et al.* (2004) have re-done Petit *et al.*-like simulations in the framework of a more self-consistent model, accounting for planetary migration. As we should have anticipated, the dynamical depletion of the Kuiper belt forces Neptune to migrate beyond 30 AU! The reason for this is that, thanks to the dynamical excitation of the distant disk provided by the embryo, 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. As shown in Figure 5, even a low mass disk of 30 $M_\oplus$ between 10 and 50 AU (7.5 $M_\oplus$ in the Kuiper belt) drives Neptune well beyond 30 AU. Stopping Neptune's migration at $\sim$ 30 AU requires a disk mass of $\sim 15 M_\oplus$ or less (depending on the initial Neptune's location). Such a mass and density profile would imply that there was only 3.75 $M_\oplus$ of material originally in the Kuiper belt between 40 and 50 AU, which is less than the mass required (10–30 $M_\oplus$) by the models of accretion of Kuiper belt bodies (Stern & Colwell 1997a; Kenyon & Luu 1999a, 1999b).

This result is quite general. We must conclude that Neptune never saw the missing mass of the Kuiper belt. Thus, any mechanism that depletes the Kuiper belt mass by handing it off to Neptune (for an other example see Nagasawa & Ida 2000) can be ruled out.

The collisional grinding scenario was proposed by Stern & Colwell (1997b) and Davis & Farinella (1997, 1998). A massive Kuiper belt with large eccentricities and inclinations would undergo a very intense collisional cascade. 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 scenario, however, has several apparent problems. First, it requires a peculiar size distribution, such that all of the missing mass was contained in small, easy to break, objects, while the number of large object was essentially identical to the current one.

Second, in order to reduce the mass of the Kuiper belt to less than an Earth mass over the age of the Solar System, Stern and Colwell (1997b) required large eccentricities and inclinations ($e \sim 0.25$ and/or $i \sim 7°$). This excitation is significantly larger than that characterizing the cold population.

Third, most of the binaries in the cold population would not survive the collisional grinding phase (Petit & Mousis 2004). In fact, the Kuiper belt binaries have large separations, so that it is easily shown that the impact on the satellite of a projective only 1% of the satellite's mass could unbind the binary. If the collisional activity was strong enough to cause an effective reduction of the overall mass of the Kuiper belt, these kind of collisions would have been extremely common. Thus, the existence of a large number of binaries is strong evidence against the collision grinding scenario.

So, at this point in the history of the field, we seemed to be faced with an unsolvable contradiction — we need to start with a large amount of mass within the 40 – 50 AU region in order to grow the Kuiper belt the we see, but we have no way to get rid of this mass and still be consistent with the observations. However, there is an implicit assumption that is being made in the above argue. In particular, we have been assuming that the cold population formed where we see it.

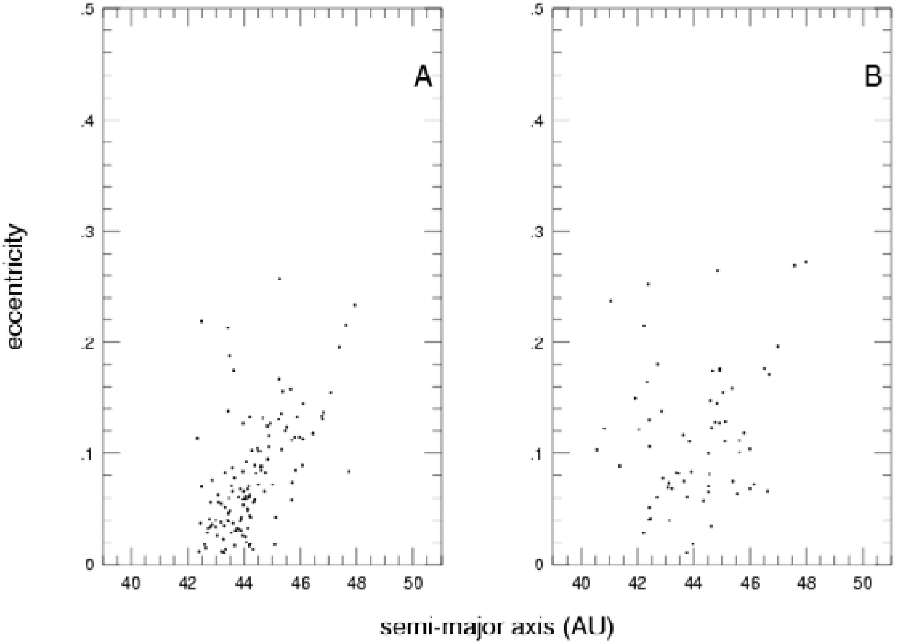

Figure 6. Left: the observed semi-major axis vs eccentricity distribution of the cold population. Only bodies with multi-opposition orbits and $i < 4°$ are taken into account. Right: the resulting orbital distribution in the scenario proposed by Levison & Morbidelli (2003).

8. Circumventing the mass depletion problem

A possible way out of this mass depletion problem has been proposed by Levison & Morbidelli (2003). In their preferred scenario, the primordial edge of the massive protoplanetary disk was somewhere around 30–35 AU and the *entire* Kuiper belt population — not only the hot component as in Gomes's scenario — formed within this limit and was transported to its current location during Neptune's migration.

In Levison & Morbidelli (2003)'s scenario, the transport process for the cold population was different from the one found by Gomes (2003) for the hot population. The bodies in the cold population were trapped in the 1:2 resonance with Neptune and transported outwards within the resonance, until they were progressively released due to the non-smoothness of the planetary migration. In the standard adiabatic migration scenario (Malhotra 1995), there would be a resulting correlation between the eccentricity and the semi-major axis of the released bodies. However Levison & Morbidelli found that this correlation is broken by a secular resonance embedded in the 1:2 mean motion resonance. Simulations of this process can match the observed (a, e) distribution of the cold population fairly well (see Figure 6), while the initially small inclinations are only very moderately perturbed.

In this scenario, the small mass of the current Kuiper belt population is simply due to the fact that, presumably, only a small fraction of the massive disk population was initially trapped in the 1:2 resonance and released on stable non-resonant orbits. The preservation of the binary objects is not a problem, because these objects were moved out

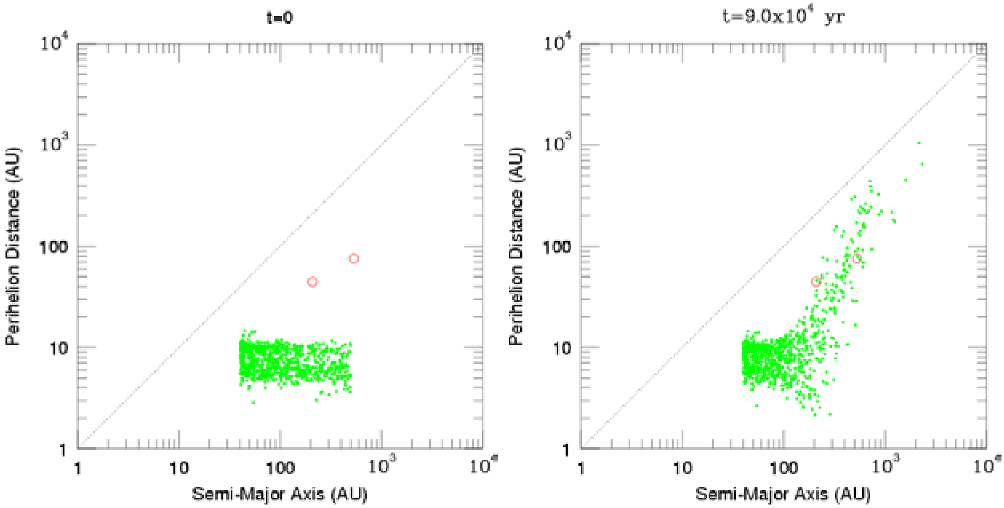

Figure 7. The semi-major axis – perihelion distance distribution of the scattered disk before and after Morbidelli & Levison (2004)'s nominal stellar passage with $q_\star = 800$ AU. The left panel is taken directly from the simulations in Dones *et al.* (2004). The right panel shows the effect of the passage. The pluses show the location of the simulated particles, while the filled circles show the real location of 2000 CR$_{105}$ and (90377) Sedna.

of the massive disk in which they formed by a gentle dynamical process. The final position of Neptune would simply reflect the primitive truncation of the protoplanetary disk (see §6). Conversely, this model opens again the problem of the origin of different physical properties of the cold and hot populations, because they would have both originated within 35 AU, although in somewhat different parts of the disk. This scenario also makes a simple prediction that will be confirmed or denied by future observations: the edge of the cold classical belt is exactly at the location of the 1:2 resonance.

9. Origin of the extended scattered disk

Explaining the origin of the orbits of 2000 CR$_{105}$ ($a \sim 230$ AU, $q \sim 45$ AU) and (90377) Sedna ($a = 509$ AU, $q = 76$ AU) is a major test for our understanding of the primordial evolution of the outer Solar System. Gladman *et al.* (2001) showed that 2000 CR$_{105}$ could not have been a normal member of the scattered disk that had its perihelion distance increased by chaotic diffusion. The same conclusion also clearly applies to (90377) Sedna. Thus, as with the Kuiper belt, these extended scattered disk objects are the result of dynamical processes that are no longer acting in the Solar System. Understanding how these objects could have formed will supply important clues to the origin of the Solar System.

In Morbidelli & Levison (2004), we explored five seemingly promising mechanisms for explaining the origin of the orbits of these peculiar objects: (i) the passage of Neptune through a high-eccentricity phase, (ii) the past existence of massive planetary embryos in the Kuiper belt or the scattered disk, (iii) the presence of a massive trans-Neptunian disk at early epochs that perturbed highly-inclined scattered disk objects, (iv) encounters with other stars that perturbed the orbits of some of the Solar System's trans-Neptunian planetesimals, and (v) the capture of extra-solar planetesimals from low mass stars or brown dwarfs encountering the Sun.

Of the five mechanisms listed above, only the two related to early stellar passages appear satisfactory. By satisfactory, we mean capable of producing the orbits of both

2000 CR_{105} and Sedna at the same time, without generating a larger population of extended scattered disk objects with $q \sim 45$ AU but $a < 200$ AU. In our analysis, we put a lot of emphasis on the absence of detections of bodies with such orbital characteristics. Since observational biases (given an object's perihelion distance and absolute magnitude, and a survey's limiting magnitude of detection) sharply favor the discovery of objects with small semi-major axes, we think that it would be unlikely that the first two discovered bodies with $q > 44$ AU had $a > 200$ AU if the real semi-major axis distribution in the extended scattered disk were skewed toward smaller a.

We believe that the most likely scenario for the origin of these objects is that they were originally normal scattered disk objects that got placed on their current orbits by a close stellar encounter early in the history of the Solar System. Figure 7 shows an example of an encounter between a young scattered disk and an $1\,M_{\odot}$ star that passes at 800 AU on a slightly hyperbolic orbit. The initial condition for the scattered disk was taken from the simulations of Dones *et al.* (2004) and represent the scattered disk at 10^5 yrs. As the figure shows, there is excellent agreement between the model and the observations in that both the orbits of 2000 CR_{105} and Sedna are reproduced while no small semi-major axis extended scattered disk objects are produced.

10. Concluding remarks

Over ten years of dedicated surveys have revealed unexpected and intriguing properties about the trans-Neptunian population. These characteristics include the existence of a large number of bodies trapped in mean motion resonances, the overall mass deficit, the large orbital eccentricities and inclinations, and the apparent existence of an outer edge at ~ 50 AU and of a correlation between inclinations, sizes and colors. Understanding how the Kuiper belt acquired all these properties would significantly constraint models of the formation of the outer planetary system and of its primordial evolution.

Up to now, a plethora of scenarios have been proposed by theoreticians. None of them can account for all the observations alone, and the solution of the Kuiper belt primordial sculpting problem probably requires a sapient combination of the proposed models. As we first suggest in Levison & Morbidelli (2003), we currently believe that the primordial planetesimal disk was truncated inside 40 AU and that the entire Kuiper belt was pushed outward by the migration of Neptune. If true, the Kuiper belt's complex structure is the result of a combination of the mechanisms presented in Malhotra (1995), Gomes (2003), and Levison & Morbidelli (2003).

However, Kuiper belt science is a rapidly evolving one. New observations change our view of the belt every year. Since the discovery of the first trans-Neptunian object 12 years ago, several review papers have been written, and all of them are already obsolete. No doubt that this will also be the fate of this chapter, but it can be hoped that the ideas presented here can continue to guide us in the direction of further understanding of what present observations of the Kuiper belt can tell us about the formation and evolution of the outer Solar System.

References

Allen, R.L., Bernstein, G.M. and Malhotra, R. 2001, *Astrophys. J.* 549, L241
Allen, R.L., Bernstein, G.M. and Malhotra, R. 2002, *Astron. J.* 124, 2949
Brown M. 2001, *Astron. J.* 121, 2804
Brown M., Trujillo C. and Rabinowitz D. 2004, *Astrophys. J. Letters*, in press.
Chiang, E.I. and Brown, M.E. 1999, *Astron. J.* 118, 1411
Davis D. R. and Farinella P. 1997. *Icarus* 125, 50

Davis D. R. and Farinella P. 1998, in: *Lunar Planet. Science Conf.* 29, 1437

Dones L., Levison, H.F., Duncan, M.J. and Weissman, P.R. 2004, *Icarus*, submitted.

Doressoundiram A., Barucci M.A., Romon J. and Veillet C. 2001, *Icarus* 154, 277

Duncan, M. J., Levison, H. F. and Budd, S. M. 1995, *Astron. J.* 110, 3073

Duncan, M. J. and Levison, H. F. 1997, *Science* 276, 1670

Emel'yanenko V. V., Asher D. J. and Bailey M. E. 2003, *Mon. Not. R. Astron. Soc.* 338, 443

Gladman, B., Kavelaars, J.J., Petit, J.M., Morbidelli, A., Holman, M.J. and Loredo, Y. 2001, *Astron. J.* 122, 1051

Gladman B., Holman M., Grav T., Kaavelars J.J., Nicholson P., Aksnes K. and Petit J.M. 2002, *Icarus* 157, 269

Gomes R. S. 2000, *Astron. J.* 120, 2695

Gomes R.S. 2003, *Icarus* 161, 404

Gomes R.S., Morbidelli A. and Levison H.F. 2004, *Icarus* 170, 492

Hahn J. M. and Malhotra R. 1999, *Astron. J.* 117, 3041

Henrard J. 1982, *Celest. Mech.* 27, 3

Hollenbach, D. and Adams, F.C. 2003, in: L. Caroff and D. Backman, (eds.) *Debris Disks and the Formation of Planets,* (San Francisco: ASP), in press

Ida S., Larwood J. and Burkert A. 2000, *Astrophys. J.* 528, 351

Jewitt, D. and Luu, J. 1993, *Nature* 362, 730

Jewitt, D., Luu, J. and Chen, J. 1996, *Astron. J.* 112, 1225

Kenyon, S.J. and Luu, J.X. 1998, *Astron. J.* 115, 2136

Kenyon, S.J. and Luu, J.X. 1999a, *Astron. J.* 118, 1101

Kenyon, S.J. and Luu, J.X. 1999b, *Astrophys. J* 526, 465

Kenyon S.J. and Bromley B.C. 2002, *Astron. J.*, 124, 1757

Kobayashi H. and Ida S. 2001, *Icarus* 153, 416

Levison H.F. and Stern S.A. 2001, *Astron. J.* 121, 1730

Levison H.F. and Morbidelli A. 2003, *Nature* 426, 419

Malhotra, R. 1993, *Nature* 365, 819

Malhotra, R. 1995, *Astron. J.* 110, 420

Malhotra, R. 1996, *Astron. J.* 111, 504

Melita M., Larwood J., Collander-Brown S, Fitzsimmons A., Williams I.P. and Brunini A. 2002, in: *Asteroids, Comets, Meteors*, ESA Spec. Publ. series, SP-500, p. 305

Morbidelli A. and Valsecchi G. B. 1997, *Icarus* 128, 464

Morbidelli A. and Brown M. 2003, in: M. Festou et al. (eds.) *Comets II*, Univ. Arizona Press, Tucson, in press

Morbidelli A. and Levison H.F. 2004, *Astron. J.*, in press.

Nagasawa M. and Ida S. 2000, *Astron. J.* 120, 3311

Petit J. M., Morbidelli A. and Valsecchi, G. B. 1999, *Icarus* 141, 367

Petit J. M. and Mousis O. 2004, *Icarus* 168, 409

Stern, S. A. 1996, *Astron. J.* 112, 1203

Stern, S. A. and Colwell, J. E. 1997a, *Astron. J.* 114, 841

Stern, S. A. and Colwell, J. E. 1997b, *Astrophys. J.* 490, 879

Stone, J.M., Gammie, C.F., Balbus, S.A. and Hawley, J.F. 1998, in: V. Mannings, A.P. Boss, and S.S. Russell (eds.) *Protostars and Planets IV*, p. 589

Tegler, S. C. and Romanishin, W. 2000, *Nature* 407, 979

Trujillo C.A. and Brown M.E. 2001, *Astrophys. J* 554, 95

Trujillo, C. A., Jewitt D. C. and Luu J. X. 2001, *Astron. J.* 122, 457

Trujillo C.A. and Brown M.E. 2002, *Astrophys. J.* 566, 125

Trujillo C.A., Brown M.E. 2003, *Earth, Moon, Planets* 92, 99

Weidenschilling S. 2003, in: M. Festou *et al.* (eds.), *Comets II*, Univ. Arizona Press, Tucson, in press

Youdin, A.N. and Shu, F.H. 2002, *Astron. J.* 580, 494