Available online at www.sciencedirect.com
SCIENCE DIRECT®
C. R. Physique 4 (2003) 809–817
COMPTES RENDUS PHYSIQUE

New frontiers in the Solar System: trans-Neptunian objects/Les nouvelles frontières du système solaire : les objets transneptuniens

The primordial sculpting of the Kuiper belt

Alessandro Morbidelli [a,*], Hal F. Levison [b]

[a] *Observatoire de la Côte d'Azur, B.P. 4229, 06304 Nice cedex 4, France*
[b] *SWRI, 1050 Walnut Street, Suite 426, Boulder, CO 80302, USA*

Presented by Pierre Encrenaz

Abstract

Understanding how the Kuiper belt acquired its puzzling present orbital structure will provide insight into the formation of the outer planetary system and on its early evolution. We outline a scenario of primordial sculpting – issued from a combination of mechanisms proposed by various authors – that seems to explain most of the observed properties of the Kuiper belt. 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 of the detailed sculpture – at least of its rough cast. ***To cite this article: A. Morbidelli, H.F. Levison, C. R. Physique 4 (2003).***

Résumé

L'évolution primitive de la ceinture de Kuiper et l'acquisition de sa structure actuelle. Comprendre comment la ceinture de Kuiper a acquit sa structure actuelle apporterait également une nouvelle compréhension de la formation et de l'évolution précoce du Système Solaire externe. Nous traçons les grandes lignes d'un scénario cohérent – issu de la combinaison de plusieurs mécanismes déjà proposés dans la littérature – qui pourrait expliquer la plupart des propriétés observées de la ceinture de Kuiper. Certains aspects ne sont pas totalement clairs, et d'autres ne sont peut-être pas totalement corrects. Mais, pour la première fois on commence à entrevoir la suite des événements qui ont donné à la ceinture de Kuiper sa forme actuelle. ***Pour citer cet article : A. Morbidelli, H.F. Levison, C. R. Physique 4 (2003).***

Keywords: Kuiper belt; orbital structure

Mots-clés : ceinture de Kuiper ; structure orbitale

1. Introduction

When Edgeworth and Kuiper conjectured the existence of a belt of small bodies beyond Neptune – the presently called Kuiper belt – they certainly were imagining a disk of planetesimals preserving the pristine conditions of the proto-planetary disk. But, since the first discoveries of trans-Neptunian objects, astronomers have realized that this picture is not correct: the disk has been affected by a number of processes which have altered its original structure. The Kuiper belt thus provides us with a large number of clues to understand what happened in the outer Solar System during the primordial ages.

In this respect, it is important to distinguish between the two different structures, called *Kuiper belt* and *scattered disk* (see Fig. 1). We call *scattered disk* the region of the orbital space that can be visited by bodies that had a close encounter with Neptune at least once during the age of the Solar System, assuming no substantial modification of the planetary orbits (Levison and Duncan, [1]; Duncan and Levison, [2]). We then call *Kuiper belt* the complement of the scattered disk in the $a > 30$ AU

* Corresponding author.
E-mail address: morby@obs-nice.fr (A. Morbidelli).

1631-0705/$ – see front matter
doi:10.1016/j.crhy.2003.10.001

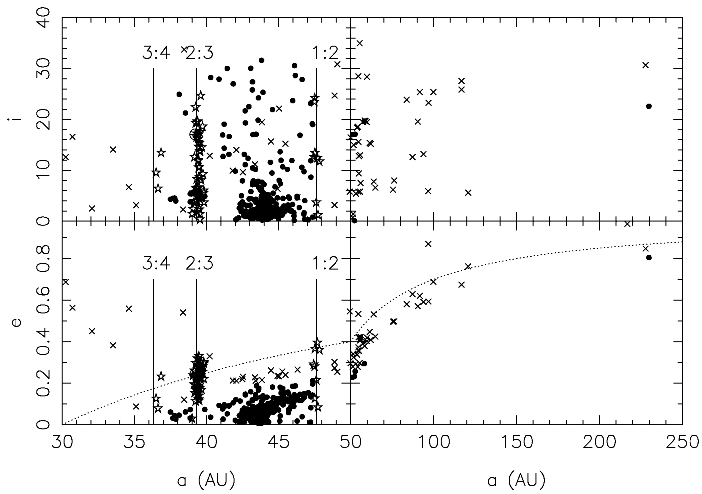

Fig. 1. The orbital distribution of multi-opposition trans-Neptunian bodies, as of 3 March 2003. Scattered disk bodies are represented as a cross, Kuiper belt bodies as stars or dots, depending if they are in a major mean motion resonance with Neptune (*resonant objects*) or not (*classical objects*). We qualify that, in absence of long term numerical integrations of the evolution of all the objects and because of the uncertainties in the orbital elements, some bodies could have been miss-classified. Thus, the figure should be considered as an indicative representation of the various subgroups that compose the trans-Neptunian population. The dotted curve denotes $q = 30$ AU. The vertical solid lines mark the locations of the 3:4, 2:3 and 1:2 mean motion resonances with Neptune. The orbit of Pluto is represented by a crossed circle.

region. The bodies that belong to the scattered disk in this classification do not provide us many direct clues to uncover the primordial architecture of the Solar System. In fact their current orbits might have been achieved starting from quasi-circular ones in Neptune's zone by pure dynamical evolution, in the framework of the current architecture of the Solar System. The opposite is true for the orbits of the Kuiper belt objects, which are essentially 'frozen' (Duncan et al., [3]). 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, [4]). Therefore, all the properties that make the current Kuiper belt very different from an accretion disk must have been acquired during a primordial phase when the Solar System looked very different from the present state. We now enumerate all these puzzling properties:

- (i) *The mass deficit*. 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. [5], Chiang and Brown [6], Trujillo et al. [7], Gladman et al. [8]). However, the primordial mass of the protoplanetary disk in the same region should have been several tens of Earth masses. This mass is in fact required for the accretion process to produce the observed objects within a reasonable time of several 10^7 to 10^8 years (Stern, [4]; Stern and Colwell, [9]; Kenyon and Luu, [10–12]).
- (ii) *The existence of resonant populations*. A glance at Fig. 1 shows that the Kuiper belt is made of two distinct sub-populations: the *resonant population* (star symbols) and the *classical belt* (dots). 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. It is well known that mean motion resonances offer a protection mechanism against close encounters with the resonant planet (Cohen and Hubbard, [13]). For this reason, the resonant population can have perihelion distances much smaller than the classical belt objects, and even Neptune-crossing orbits ($q < 30$ AU) as in the case of Pluto (Malhotra, [14]). The bodies in the 2:3 resonance are often called *Plutinos*, for the analogy of their orbit with that of Pluto. According to Trujillo et al. [15] the resonant populations account altogether for 10% of the total Kuiper belt population.
- (iii) *The bi-modal inclination distribution of the classical belt*. Fig. 1 shows that the inclination of the bodies in the Kuiper belt can be very large. An analysis accounting for observational biases made by Brown [16] suggests that the real inclination distribution of the classical belt is bi-modal. About half of the objects have an inclination smaller than 4°, while the remaining half have a very distended 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 and Stern [17]; Trujillo and Brown, [18]); moreover there is a statistically significant difference between the color distributions within the two populations (Tegler and Romanishin, [19]; Doressoundiram et al., [20]; Trujillo and Brown, [21]).

(iv) The outer edge of the classical belt. It was generally expected that the mass of the Kuiper belt should smoothly decrease with heliocentric distance – or perhaps even increase in number density by a factor of ∼100 back to the level of the extrapolation of the minimum mass Solar nebula beyond the region of Neptune's influence (Stern, [4]). However the persisting lack of detection of objects beyond about 50 AU (Allen et al., [22,23]) cannot be explained by observational biases, but implies a statistically significant steep drop off in number density of large objects (Trujillo and Brown, [7]). Fig. 1 suggests that the edge of the classical belt coincides with the location of the 1:2 mean motion resonance with Neptune.

(v) The odd (a, e) shape of the classical belt. Fig. 1 shows that there are many objects with $e \sim 0$ interior to 45 AU, but beyond $a = 45$ AU e on average increases with a. Since the lower boundary of this (a, e) distribution beyond 45 AU does not follow a curve of constant perihelion, this is most likely not the result of an observational bias. Curiously, both the cold and the hot populations have similar distributions on the (a, e) plane.

(vi) Existence of an 'extended scattered disk' population. Fig. 1 shows the existence of bodies with $a > 50$ AU and highly eccentric orbits that do not belong to the scattered disk. 5 objects are currently known, including 2000 CR_{105} ($a = 230$ AU, perihelion distance $q = 44.17$ AU and inclination $i = 22.7^{\circ}$), but our classification is uncertain for the reasons explained in the figure caption. We call these objects *extended scattered disk* objects for two reasons: (a) they do not belong to the scattered disk according to our definition but are very close to its boundary; and (b) a body of ~ 300 km like 2000 CR_{105} presumably formed much closer to the Sun, where the accretion timescale was sufficiently short (Stern, [4]), implying that it has been subsequently transported in semi-major axis until its current location was reached. This hypothesis suggests that in the past the true scattered disk extended well beyond its present boundary in perihelion distance. Given that the observational biases become rapidly more severe with increasing perihelion distance and semi-major axis, the currently known extended scattered disk objects may be like the tip of an iceberg, e.g., the emerging representatives of a conspicuous population (Gladman et al., [24]).

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 in the primordial sculpting of the trans-Neptunian population. A more exhaustive review can be found in Morbidelli and Brown [25].

2. Origin of the resonant populations

Fernández and Ip [26] showed that, while scattering primordial planetesimals, Neptune should have migrated outwards. Malhotra [27,28] 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, [29]), some of the Kuiper belt objects swept by a mean motion resonance would have been captured into resonance; they would have subsequently followed the resonance in its migration, while increasing their eccentricity. 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 Fig. 2). Reproducing the observed range of eccentricities of the resonant bodies requires that Neptune migrated by 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 process have been investigated in detail by Gomes [30], who concluded that, although large inclinations can be achieved, the resulting proportion between the number of high inclination versus low inclination bodies and their distribution in the eccentricity vs. inclination plane do not reproduce well the observations. According to Gomes [31] 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.

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 and Malhotra [32], who simulated Neptune's migration using a disk of Lunar to Martian-mass planetesimals, did not obtain any permanent capture.

3. Origin of the hot population

Gomes [31] showed that the Neptune migration, in addition to the resonant populations, can also explain the origin of the hot population. Like Hahn and Malhotra [32] Gomes simulated Neptune's migration, starting from about 15 AU, by the interaction with a massive planetesimal disk extending from beyond Neptune's initial position. But, taking advantage of the

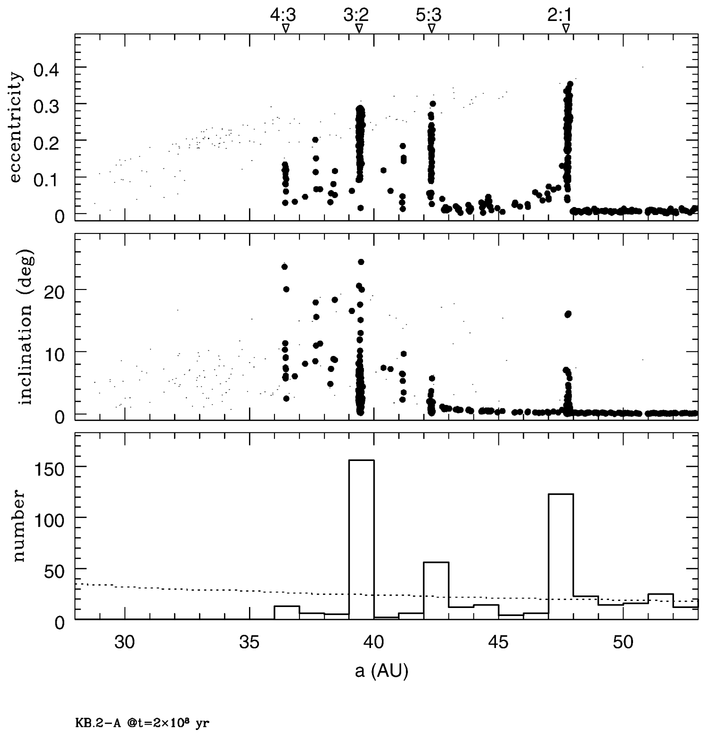

Fig. 2. Final distribution of the Kuiper belt bodies according to the sweeping resonances scenario (courtesy of R. Malhotra). The simulation is done by numerical integrating, over a 200 Myr timespan, the evolution of 800 test particles on initial quasi-circular and coplanar orbits. The planets are forced 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. In the lowest panel, the solid line is the histogram of semi-major axis of the 'surviving' particles; the dotted line is the initial distribution.

improved computer technology, he used 10 000 particles to simulate the disk population, with individual masses roughly equal to twice the Pluto's mass, while Hahn and Malhotra used only 1000 particles, with Lunar to Martian masses. In his simulations, during its migration Neptune scattered the planetesimals and formed 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 eccentricity would have eventually increased back to Neptune-crossing values, the dynamics being reversible. But Neptune's migration broke the reversibility, and some of the decoupled bodies managed to escape from the resonances, and remained permanently trapped in the Kuiper belt. As shown in Fig. 3, the current Kuiper belt would therefore be the result of the superposition of these bodies with the local population, originally formed beyond 30 AU, which stays dynamically cold because only moderately excited (by the resonance sweeping mechanism, as in Fig. 2).

The migration mechanism is sufficiently slow (several 10^7 y) that the scattered particles have the time to acquire very large inclinations, consistent with the observed hot population. The resulting inclination distribution of the bodies in the classical belt is bimodal, and quantitatively reproduces the de-biased inclination distribution computed by Brown [16] from the observations.

In Gomes [31] simulations an extended scattered disk is also formed beyon 50 AU. Although bodies on orbits similar to that of 2000 CR_{105} are not obtained in the nominal simulations, other tests done in [31] and new simulations (Gomes, personal communication) are suggestive that such orbits could be achieved in the framework of the same scenario.

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 $\sim$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 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.

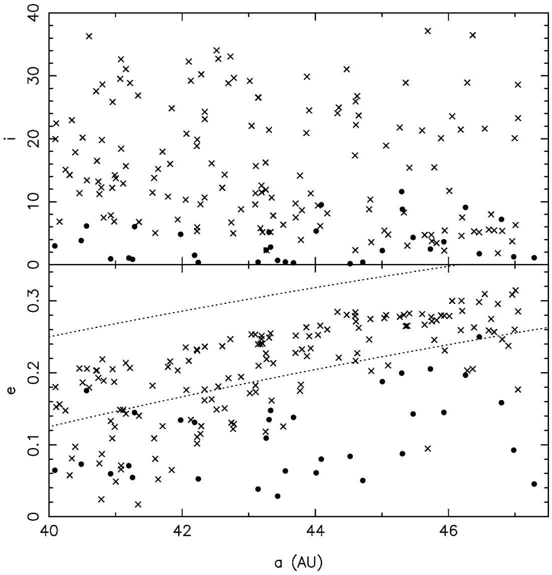

Fig. 3. The orbital distribution in the classical belt according to Gomes' simulations. The dots denote the local population, which is only moderately dynamically excited. The crosses denote the bodies that were originally inside 30 AU. Therefore, the resulting Kuiper belt population is the superposition of a dynamically cold population and of a dynamically hot population, which gives a bi-modal inclination distribution comparable to that observed. The dotted curves in the eccentricity vs. semi-major axis plot correspond to $q = 30$ AU and $q = 35$ AU. Courtesy of R. Gomes.

4. Origin of the outer edge of the Kuiper belt

The existence of an outer edge of the Kuiper belt is a very intriguing property. At least three mechanisms for its origin have been proposed, none of which has raised the general consensus of the community of the experts.

Brunini and Melita [33] showed with numerical simulations that a Martian mass body residing for 1 Gy on an orbit with $a \sim 60$ AU and $e \sim 0.15$–0.2 could have scattered into Neptune-crossing orbits most of the Kuiper belt bodies originally in the 50–70 AU range, leaving this region strongly depleted and dynamically excited. The apparent edge at 50 AU might be simply the inner edge of a similar gap in the distribution of Kuiper belt bodies. A problem of the Brunini and Melita scenario is that there are no evident dynamical mechanisms that would ensure the later removal of the massive body from the system. In other words, the massive body should still be present, somewhere in the ∼50–70 AU region. A Mars-size body with 4% albedo at 70 AU would have apparent magnitude brighter than 20, so that, if its inclination is small ($i < 10°$) it is unlikely that it escaped detection in the numerous wide field ecliptic surveys that have been performed up to now, and in particular in that led by Trujillo and Brown [18]. We remark that a small inclination should be expected if the putative Matian body was originally a scattered disk object whose eccentricity (and inclination) were damped by dynamical friction (as conjectured by Brunini and Melita) or if it reached its required heliocentric distance by migrating through the primordially massive Kuiper belt (the most likely evolution according to Gomes et al., [34]).

Weidenshilling [35] suggested that the outer edge of the Kuiper belt is the result of the facts that accretion takes longer with increasing heliocentric distance and small planetesimals drift inwards due to gas drag. This leads to a steepening of the radial surface density gradient of solids. The edge effect is augmented because, at whatever distance large bodies can form, they capture the ∼m-sized bodies spiraling in from further out. The net result of the process, as shown by Weidenschilling's numerical modeling, is production of an edge, where both the surface density of solid matter and the mean size of planetesimals decrease sharply with distance.

A third possibility is that the planetesimal disk was truncated by the passage of a star in the vicinity of the Sun. Ida et al. [36] and Kobayashi and Ida [37] showed that the resulting eccentricities and inclinations of the planetesimals would depend critically on a/D, where a is their semi-major axis and D is the heliocentric distance of the stellar encounter. A stellar encounter at ∼200 AU would make most of the bodies beyond 50 AU so eccentric to intersect the orbit of Neptune, which eventually would produce the observed edge (Melita et al., [38]). An interesting constraint on the time at which such an encounter occurred is set by the existence of the Oort cloud. Levison et al. [39] showed that the encounter had to occur much earlier than ∼10 My after the formation of Uranus and Neptune, otherwise most of the existing Oort cloud would have been ejected to interstellar space and many of the planetesimals in the scattered disk would have had their perihelion distance lifted beyond Neptune, decoupling from the planet. As a consequence, the extended scattered disk population, with $a > 50$ AU and $40 < q < 50$ AU, would have had a mass comparable or larger than that of the resulting Oort cloud, hardly compatible with the few detections of extended scattered disk objects performed up to now. An encounter with a star during the first million year from planetary formation is a likely event if the Sun formed in a stellar cluster (Bate et al., [40]). At such an early time, presumably the Kuiper belt objects were not yet fully formed (Stern, [4], Kenyon and Luu, [10]). In this case, the edge of the belt would be at a heliocentric distance corresponding to a post-encounter eccentricity excitation of ∼0.05, a threshold value below which collisional damping is efficient and accretion can recover, and beyond which the objects rapidly grind down to dust (Kenyon and Bromley, [41]).

None of the three presented scenario can explain why the outer edge of the Kuiper belt lies approximately at the location of the 1:2 resonance.

The origin of the orbit of 2000 CR_{105} can also be attributed to a stellar encounter, but not to the one forming the edge of the Kuiper belt. In fact, such a close encounter would produce also a relative overabundance of bodies with perihelion distance similar to that of 2000 CR_{105} but with semi-major axis in the 50–200 AU range. These bodies have never been discovered despite of the more favorable observational biases. In order that only bodies with $a > 200$ AU have their perihelion distance lifted, a second stellar passage at about 800 AU is required (Morbidelli and Levison, [42]).

5. The mass deficit of the cold population

The scenario proposed by Gomes [31] reduces the problem of the mass depletion of the Kuiper belt to the sole cold population. In fact, in Gomes' 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, which naturally explains the current low mass of the hot population. However, the population originally in the 40–50 AU range – which would constitute the cold population in Gomes scenario – should have been only moderately excited and not dynamically depleted, so that it should have preserved most of its primordial mass.

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 region and the collisional comminution of most of the mass of the Kuiper belt into dust.

The dynamical depletion mechanism was proposed by Morbidelli and Valsecchi [43] and Petit et al. [44]. 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 $\sim 10^8$ 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. [44] 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. [31] have re-done a Petit et al.-like simulations in the framework of a more self-consistent model accounting for planetary migration. As expected, the dynamical depletion of the Kuiper belt largely enhances Neptune's migration. 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 Fig. 4 even a low mass disk of $30M_\oplus$ between 10 and 50 AU ($7.5M_\oplus$ in the Kuiper belt) drives Neptune well beyond 30 AU. Halting Neptune's migration at ∼30 AU requires a disk mass of $\sim 15M_\oplus$ or less (depending on the initial Neptune's location). Such a mass and density profile would imply only $3.75M_\oplus$ of material originally in the Kuiper belt between 40 and 50 AU, which is less than the mass required (10–$30M_\oplus$) by the models of accretion of Kuiper belt bodies (Stern and Colwell, [9]; Kenyon and Luu, [11,12]).

A priori, for what concerns Neptune's migration, there is no evident difference between the case where the Kuiper belt is excited to Neptune-crossing orbit by a planetary embryo or by some other mechanism, such as the primordial secular resonance sweeping (Nagasawa and Ida, [45]). Therefore, we conclude that Neptune never saw the missing mass of the Kuiper belt.

The collisional grinding scenario was proposed by Stern and Colwell [46] and Davis and Farinella [47,48]. 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 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 [46] required a large eccentricity and inclination excitation ($e \sim 0.25$ and/or $i \sim 7^\circ$). This excitation is significantly larger than that characterizing the cold population.

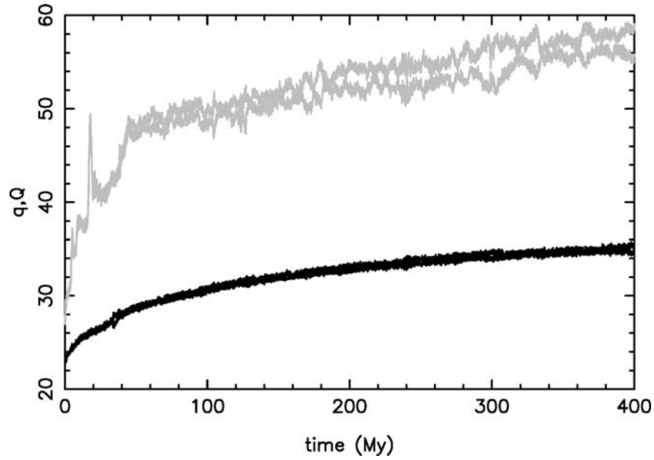

Fig. 4. A self-consistent simulation of the Petit et al. (1999) scenario for the excitation and dynamical depletion of the Kuiper belt (from Gomes et al., 2003). Neptune is originally assumed at ~ 23 AU and an Earth-mass embryo at ~ 27 AU. Both planets are embedded in a $30M_\oplus$ disk, extending from 10 to 50 AU with a r^{-1} surface density profile ($7.5M_\oplus$ between 40 and 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, 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).

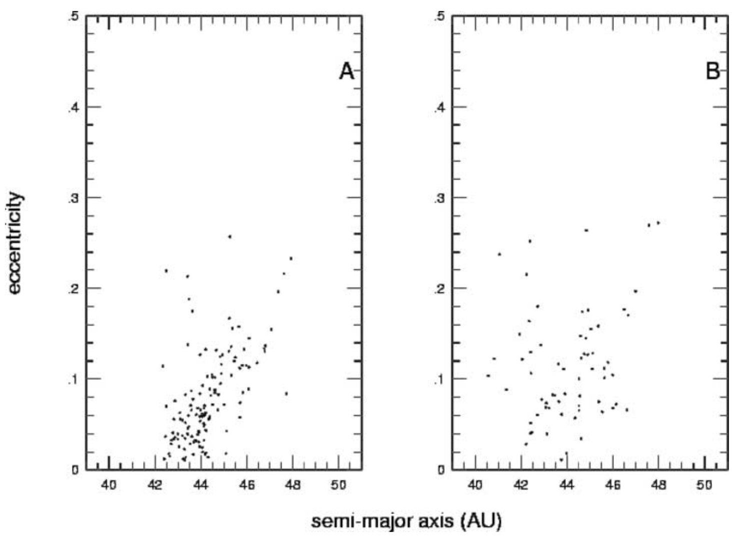

Fig. 5. Left: the observed semi-major axis vs eccentricity distribution of the cold population. Only bodies with multi-opposition orbits and $i < 4^\circ$ are taken into account. Right: the resulting orbital distribution in the scenario proposed by Levison and Morbidelli [50].

Third, most of the binaries in the cold population would not survive the collisional grinding phase (Petit and Mousis, [49]). In fact, the Kuiper belt binaries have large separations, so that it can be easily computed that the impact on the satellite of a 100 times less massive projectile with a speed of 1 km/s would give the former an impulse velocity sufficient to escape to an unbound orbit. If the collisional activity was strong enough to cause an effective reduction of the overall mass of the Kuiper belt these kind of collisions had to be extremely common, so that we would not expect a significant fraction of widely separated binary objects in the current remaining population.

A possible way out of this mass depletion problem has been recently proposed by Levison and Morbidelli [50]. 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. The transport process of the cold population was different from the one found by Gomes [31] for the hot population. These bodies 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, [28]) there would be a resulting correlation between the eccentricity and the semi-major axis of the released bodies. However, this correlation is broken by a secular resonance embedded in the 1:2 mean motion resonance. Simulations of this process allow to match the observed (a, e) distribution of the cold population fairly well (see Fig. 5), while the initially small inclinations are only very moderately perturbed. In this scenario, the small mass of the current Kuiper belt population is simply the 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 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. 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. But this scenario does 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.

6. Conclusions and perspectives

Ten years of dedicated surveys have revealed unexpected and intriguing properties of the trans-Neptunian population, such as the existence of a large number of bodies trapped in mean motion resonances, the overall mass deficit, the large orbital eccentricities and inclinations, the apparent existence of an outer edge at ~ 50 AU and of a correlation among inclinations, sizes and colors. Understanding how the Kuiper belt acquired all these properties would probably constraint several aspects of the formation of the outer planetary system and of its primordial evolution.

Up to now, a portfolio 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. The Malhotra–Gomes scenario on the effects of planetary migration does a quite good job at reproducing the observed orbital distribution inside 50 AU. The apparent edge of the belt at 50 AU might be explained by a very early stellar encounter at ~ 150–200 AU. The origin of the peculiar orbit of 2000 CR_{105} could be due to a later stellar encounter at ~ 800 AU.

The most mysterious feature that remains unexplained in this combination of scenarios is the mass deficit of the cold classical belt. This suggests the possibility, proposed by Levison and Morbidelli [50] that the primordial planetesimal disk was truncated inside 40 AU and that also the cold population was pushed out from within this edge, during Neptune's migration.

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 10 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

[1] H.F. Levison, M.J. Duncan, From the Kuiper belt to Jupiter-family comets: the spatial distribution of ecliptic comets, Icarus 127 (1997) 13–32.
[2] M.J. Duncan, H.F. Levison, Scattered comet disk and the origin of Jupiter family comets, Science 276 (1997) 1670–1672.
[3] M.J. Duncan, H.F. Levison, S.M. Budd, The long-term stability of orbits in the Kuiper belt, Astron. J. 110 (1995) 3073–3083.
[4] S.A. Stern, On the collisional environment, accretion time scales, and architecture of the massive, primordial Kuiper belt, Astron. J. 112 (1996) 1203–1210.
[5] D. Jewitt, J. Luu, J. Chen, The Mauna–Kea–Cerro–Totlolo (MKCT) Kuiper belt and Centaur survey, Astron. J. 112 (1996) 1225–1232.
[6] E.I. Chiang, M.E. Brown, Keck pencil-beam survey for faint Kuiper belt objects, Astron. J. 118 (1999) 1411–1422.
[7] C.A. Trujillo, M.E. Brown, The radial distribution of the Kuiper belt, Astroph. J. 554 (2001) 95–98.
[8] B. Gladman, J.J. Kavelaars, J.M. Petit, A. Morbidelli, M.J. Holman, Y. Loredo, The structure of the Kuiper belt: Size distribution and radial extent, Astron. J. 122 (2001) 1051–1066.
[9] S.A. Stern, J.E. Colwell, Accretion in the Edgeworth–Kuiper belt: Forming 100–1000 Km radius bodies at 30 AU and beyond, Astron. J. 114 (1997) 841–849.
[10] S.J. Kenyon, J.X. Luu, Accretion in the early Kuiper belt: I. Coagulation and velocity evolution, Astron. J. 115 (1998) 2136–2160.
[11] S.J. Kenyon, J.X. Luu, Accretion in the early Kuiper belt: II. Fragmentation, Astron. J. 118 (1999) 1101–1119.
[12] S.J. Kenyon, J.X. Luu, Accretion in the early outer Solar system, Astrophys. J. 526 (1999) 465–470.
[13] C.J. Cohen, E.C. Hubbard, The orbit of Pluto, The Observatory 85 (1965) 43–44.
[14] R. Malhotra, The phase space structure near Neptune resonances in the Kuiper belt, Astron. J. 111 (1996) 504–512.

[15] C.A. Trujillo, D.C. Jewitt, J.X. Luu, Properties of the trans-Neptunian belt: Statistics from the Canada–France–Hawaii telescope survey, Astron. J. 122 (2001) 457–473.
[16] M. Brown, The inclination distribution of the Kuiper belt, Astron. J. 121 (2001) 2804–2814.
[17] H.F. Levison, S.A. Stern, On the size dependence of the inclination distribution of the main Kuiper belt, Astron. J. 121 (2001) 1730–1735.
[18] C.A. Trujillo, M.E. Brown, The Caltech wide area sky survey: beyond (50000) quaoar, in: Proceedings of the First Decadal Review of the Edgeworth–Kuiper Belt Meeting in Antofagasta Chile, Earth Moon and Planets, 2003, in press.
[19] S.C. Tegler, W. Romanishin, Extremely red Kuiper-belt objects in near-circular orbits beyond 40 AU, Nature 407 (2000) 979–981.
[20] A. Doressoundiram, M.A. Barucci, J. Romon, C. Veillet, Multicolor photometry of trans-Neptunian objects, Icarus 154 (2001) 277–286.
[21] C.A. Trujillo, M.E. Brown, A correlation between inclination and color in the classical Kuiper belt, Astrophys. J. 566 (2002) 125–128.
[22] R.L. Allen, G.M. Bernstein, R. Malhotra, The edge of the Solar system, Astrophys. J. 549 (2001) L241–L244.
[23] R.L. Allen, G.M. Bernstein, R. Malhotra, Observational limits on a distant cold Kuiper belt, Astron. J. 124 (2002) 2949–2954.
[24] B. Gladman, M. Holman, T. Grav, J.J. Kaavelars, P. Nicholson, K. Aksnes, J.M. Petit, Evidence for an extended scattered disk, Icarus 157 (2002) 269–279.
[25] A. Morbidelli, M. Brown, The Kuiper belt and the primordial evolution of the Solar system, in: Festou, et al. (Eds.), Comet II, University Arizona Press, Tucson, AZ, 2003, in press.
[26] J.A. Fernández, W.H. Ip, Orbital expansion and resonant trapping during the late accretion stages of the outer planets, Pl. Sp. Sci. 44 (1996) 431–439.
[27] R. Malhotra, The origin of Pluto's peculiar orbit, Nature 365 (1993) 819–821.
[28] R. Malhotra, The origin of Pluto's orbit: Implications for the Solar system beyond Neptune, Astron. J. 110 (1995) 420–432.
[29] J. Henrard, Capture into resonance – An extension of the use of adiabatic invariants, Cel. Mech. 27 (1982) 3–22.
[30] R.S. Gomes, Planetary migration and Plutino orbital inclinations, Astron. J. 120 (2000) 2695–2707.
[31] R.S. Gomes, The origin of the Kuiper belt high inclination population, Icarus (2003), in press.
[32] J.M. Hahn, R. Malhotra, Orbital evolution of planets embedded in a planetesimal disk, Astron. J. 117 (1999) 3041–3053.
[33] A. Brunini, M. Melita, The existence of a planet beyond 50 AU and the orbital distribution of the classical Edgeworth–Kuiper belt objects, Icarus (2002), in press.
[34] R.S. Gomes, A. Morbidelli, H.F. Levison, Planetary migration in a planetesimal disk: why did Neptune stop at 30 AU?, 2003, in press.
[35] S. Weidenschilling, Formation of planetesimals/cometesimals in the Solar nebula, in: Festou, et al. (Eds.), Comet II, University Arizona Press, Tucson, AZ, 2003.
[36] S. Ida, J. Larwood, A. Burkert, Evidence for early stellar encounters in the orbital distribution of Edgeworth–Kuiper belt objects, Astroph. J. 528 (2000) 351–356.
[37] H. Kobayashi, S. Ida, The effects of a stellar encounter on a planetesimal disk, Icarus 153 (2001) 416–429.
[38] M. Melita, J. Larwood, S. Collander-Brown, A. Fitzsimmons, I.P. Williams, A. Brunini, The edge of the Edgeworth–Kuiper belt: Stellar encounter, trans-Plutonian planet or outer limit of the primordial Solar nebula?, in: Asteroid, Comet, Meteors, in: ESA Spec. Publ. Ser., 2002, pp. 305–308.
[39] H.F. Levison, A. Morbidelli, L. Dones, Forming the outer edge of the Kuiper belt by a stellar encounter: Constrints from the oort cloud, 2003, in preparation.
[40] M.R. Bate, I.A. Bonnell, V. Bromm, The formation of a star cluster: Predicting the properties of stars and brown dwarfs, MNRAS 339 (2003) 577–599.
[41] S.J. Kenyon, B.C. Bromley, Collisional cascades in planetesimal disks. I. Stellar flybys, Astron. J. 2002 (2002) 1757–1775.
[42] A. Morbidelli, H.F. Levison, Scenarios for the origin of an extended trans-Neptunian scattered disk, 2003, in preparation.
[43] A. Morbidelli, G.B. Valsecchi, Neptune scattered planetesimals could have sculpted the primordial Edgeworth–Kuiper belt, Icarus 128 (1997) 464–468.
[44] J.M. Petit, A. Morbidelli, G.B. Valsecchi, Large scattered planetesimals and the excitation of the small body belts, Icarus 141 (1999) 367–387.
[45] M. Nagasawa, S. Ida, Sweeping secular resonances in the Kuiper belt caused by depletion of the Solar nebula, Astron. J. 120 (2000) 3311–3322.
[46] S.A. Stern, J.E. Colwell, Collisional erosion in the primordial Edgeworth–Kuiper belt and the generation of the 30–50 AU Kuiper gap, Astroph. J. 490 (1997) 879–885.
[47] D.R. Davis, P. Farinella, Collisional evolution of Edgeworth–Kuiper belt objects, Icarus 125 (1997) 50–60.
[48] D.R. Davis, P. Farinella, Collisional erosion of a massive Edgeworth–Kuiper belt: Constraints on the initial population, Lunar Planet. Science Conf. 29 (1998) 1437–1438.
[49] J.M. Petit, O. Mousis, KBO binaries: are they really primordial? Icarus, 2003, submitted.
[50] H.F. Levison, A. Morbidelli, Pushing out the Kuiper belt, Nature, 2003, in press.