A Brief History of Transneptunian Space

Eugene Chiang, Yoram Lithwick, and Ruth Murray-Clay
University of California at Berkeley

Marc Buie and Will Grundy
Lowell Observatory

Matthew Holman
Harvard-Smithsonian Center for Astrophysics

The Edgeworth-Kuiper belt encodes the dynamical history of the outer solar system. Kuiper belt objects (KBOs) bear witness to coagulation physics, the evolution of planetary orbits, and external perturbations from the solar neighborhood. We critically review the present-day belt's observed properties and the theories designed to explain them. Theories are organized according to a possible timeline of events. In chronological order, epochs described include (1) coagulation of KBOs in a dynamically cold disk, (2) formation of binary KBOs by fragmentary collisions and gravitational captures, (3) stirring of KBOs by Neptune-mass planets ("oligarchs"), (4) eviction of excess oligarchs, (5) continued stirring of KBOs by remaining planets whose orbits circularize by dynamical friction, (6) planetary migration and capture of resonant KBOs, (7) creation of the inner Oort cloud by passing stars in an open stellar cluster, and (8) collisional comminution of the smallest KBOs. Recent work underscores how small, collisional, primordial planetesimals having low velocity dispersion permit the rapid assembly of ~5 Neptune-mass oligarchs at distances of 15–25 AU. We explore the consequences of such a picture. We propose that Neptune-mass planets whose orbits cross into the Kuiper belt for up to ~20 m.y. help generate the high-perihelion members of the hot classical disk and scattered belt. By contrast, raising perihelia by sweeping secular resonances during Neptune's migration might fill these reservoirs too inefficiently when account is made of how little primordial mass might reside in bodies having sizes on the order of 100 km. These and other frontier issues in transneptunian space are discussed quantitatively.

1. INTRODUCTION

The discovery by *Jewitt and Luu* (1993) of what many now regard as the third Kuiper belt object opened a new frontier in planetary astrophysics: the direct study of transneptunian space, that great expanse extending beyond the orbit of the last known giant planet in our solar system. This space is strewn with icy, rocky bodies having diameters ranging up to a few thousand kilometers and occupying orbits of a formerly unimagined variety.

Kuiper belt objects (KBOs) afford insight into processes that form and shape planetary systems. In contrast to main-belt asteroids, the largest KBOs today have lifetimes against collisional disruption that well exceed the age of the universe. Therefore their size spectrum may preserve a record, unweathered by erosive collisions, of the process by which planetesimals and planets coagulated. At the same time, KBOs can be considered test particles whose trajectories have been evolving for billions of years in a time-dependent gravitational potential. They provide intimate testimony of how the giant planets — and perhaps even planets that once resided within our system but have since been ejected — had their orbits sculpted. The richness of structure revealed by studies of our homegrown debris disk is unmatched by more distant, extrasolar analogs.

Section 2 summarizes observed properties of the Kuiper belt. Some of the data and analyses concerning orbital elements and spectral properties of KBOs are new and have not been published elsewhere. Section 3 is devoted to theoretical interpretation. Topics are treated in order of a possible chronology of events in the outer solar system. Parts of the story that remain missing or that are contradictory are identified. Section 4 recapitulates a few of the bigger puzzles.

Our review is packed with simple and hopefully illuminating order-of-magnitude calculations that readers are encouraged to reproduce or challenge. Some of these confirm claims made in the literature that would otherwise find no support apart from numerical simulations. Many estimates are new, concerning all the ways in which Neptune-sized planets might have dynamically excited the Kuiper belt. While we outline many derivations, space limitations prevent us from spelling out all details. For additional guid-

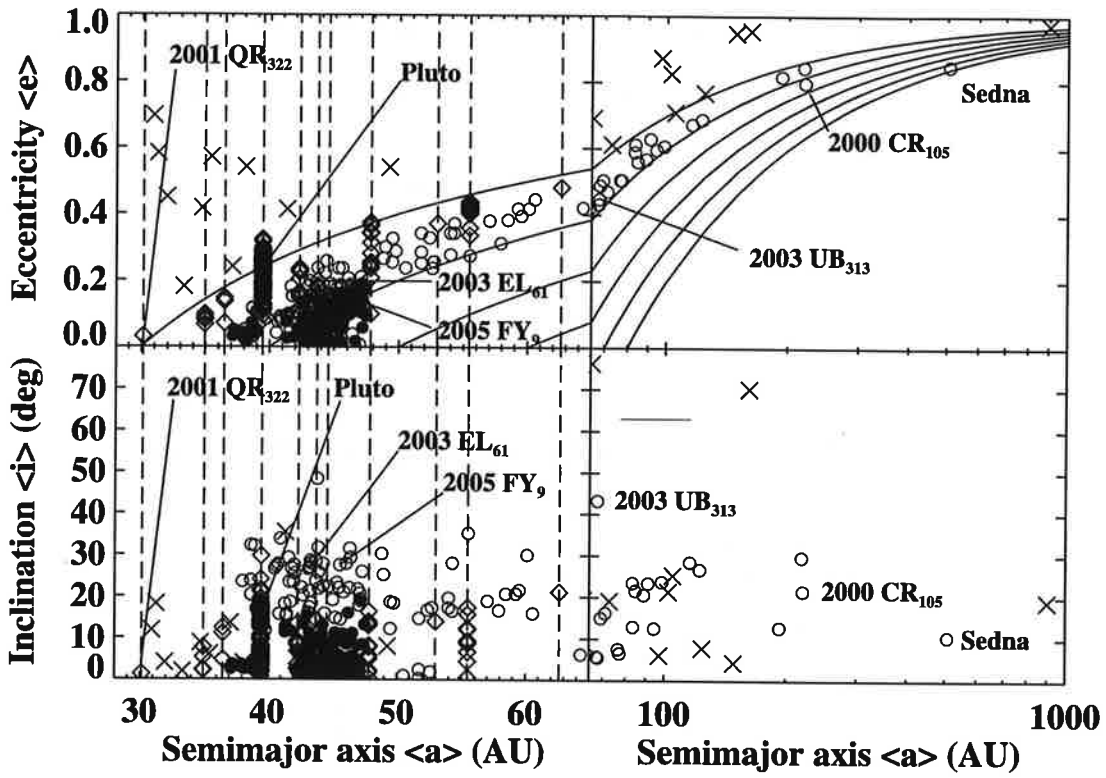

Fig. 1. Orbital elements, time-averaged over 10 m.y., of 529 securely classified outer solar system objects as of October 8, 2005. Symbols represent dynamical classes: Centaurs (×), resonant KBOs (◇), classical KBOs (●), and scattered KBOs (○). Dashed vertical lines indicate occupied mean-motion resonances; in order of increasing heliocentric distance, these include the 1:1, 5:4, 4:3, 3:2, 5:3, 7:4, 9:5, 2:1, 7:3, 5:2, and 3:1 (see Table 1). Solid curves trace loci of constant perihelion q = a(1 – e). Especially large [2003 UB_{313}, Pluto, 2003 EL_{61}, 2005 FY_9 (*Brown et al.,* 2005a,b)] and dynamically unusual [2001 QR_{322} (Trojan) (*Chiang et al.,* 2003; *Chiang and Lithwick,* 2005), 2000 CR_{105} (high q) (*Millis et al.,* 2002; *Gladman et al.,* 2002), Sedna (high q) (*Brown et al.,* 2004)] KBOs are labeled. For a zoomed-in view, see Fig. 2.

ance, see the pedagogical review of planet formation by *Goldreich et al.* (2004a, hereafter *G04*), from which our work draws liberally.

2. THE KUIPER BELT OBSERVED TODAY

2.1. Dynamical Classes

Outer solar system objects divide into dynamical classes based on how their trajectories evolve. Figure 1 displays orbital elements, time-averaged over 10 m.y. in a numerical orbit integration that accounts for the masses of the four giant planets, of 529 objects. Dynamical classifications of these objects are secure according to criteria developed by the Deep Ecliptic Survey (DES) team (*Elliot et al.,* 2005, hereafter *E05*). Figure 2 provides a close-up view of a portion of the Kuiper belt. We distinguish four classes:

1. *Resonant KBOs* (122/529) exhibit one or more mean-motion commensurabilities with Neptune, as judged by steady libration of the appropriate resonance angle(s) (*Chiang et al.,* 2003, hereafter *C03*). Resonances most heavily populated include the exterior 3:2 (Plutino), 2:1, and 5:2 (see Table 1). Of special interest is the first discovered Neptune Trojan (1:1). All resonant KBOs (except the Trojan) are found to occupy e-type resonances; the ability of an e-type resonance to retain a KBO tends to increase with the KBO's eccentricity e (e.g., *Murray and Dermott,* 1999). Unless otherwise stated, orbital elements are heliocentric and referred to the invariable plane. Several (9/122) also inhabit inclination-type (i^2) or mixed-type (ei^2) resonances. None inhabits an e_N-type resonance whose stability depends on the (small) eccentricity of Neptune. The latter observation is consistent with numerical experiments that suggest e_N-type resonances are rendered unstable by adjacent e-type resonances.

2. *Centaurs* (55/529) are nonresonant objects whose perihelia penetrate inside the orbit of Neptune. Most Centaurs cross the Hill sphere of a planet within 10 m.y. Centaurs are likely descendants of the other three classes, recently dislodged from the Kuiper belt by planetary per-

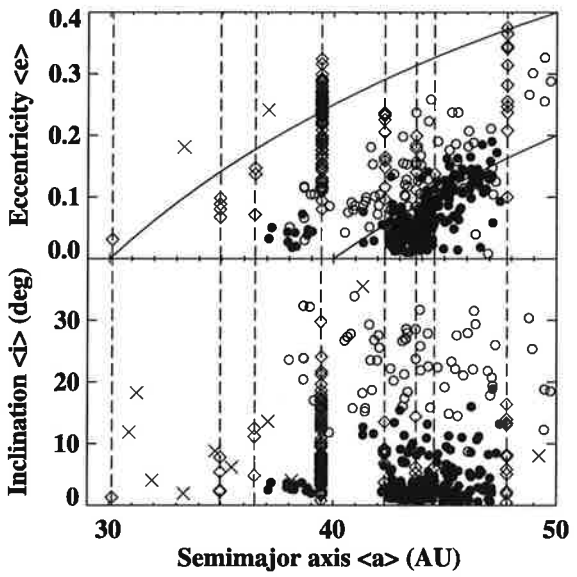

Fig. 2. Same as Fig. 1, zoomed in.

TABLE 1. Observed populations of Neptune resonances (securely identified by the DES team as of October 8, 2005).

Order 0		Order 1		Order 2		Order 3		Order 4	
m:n	#	m:n	#	m:n	#	m:n	#	m:n	#
1:1	1	5:4	4	5:3	9	7:4	8	9:5	2
		4:3	3	3:1	1	5:2	10	7:3	1
		3:2	72						
		2:1	11						

turbations (*Holman and Wisdom,* 1993, hereafter *HW93*; *Tiscareno and Malhotra,* 2003). They will not be discussed further.

3. *Classical KBOs* (246/529) are nonresonant, nonplanet-crossing objects whose time-averaged $\langle e \rangle \leq 0.2$ and whose time-averaged Tisserand parameters

$$\langle T \rangle = \langle (a_N/a) + 2\sqrt{(a/a_N)(1 - e^2)} \cos \Delta i \rangle \qquad (1)$$

exceed 3. Here Δi is the mutual inclination between the orbit planes of Neptune and the KBO, a is the semimajor axis of the KBO, and a_N is the semimajor axis of Neptune. In the circular, restricted, three-body problem, test particles with $T > 3$ and $a > a_N$ cannot cross the orbit of the planet [i.e., their perihelia $q = a(1 - e)$ remain greater than a_N]. Thus, classical KBOs can be argued to have never undergone close encounters with Neptune in its current nearly circular orbit and to be relatively pristine dynamically. Indeed, many classical KBOs as identified by our scheme have low inclinations $\langle i \rangle < 5°$ ("cold classicals"), though some do not ("hot classicals"). Our defining threshold for $\langle e \rangle$ is arbitrary; like our threshold for $\langle T \rangle$, it is imposed to suggest — perhaps incorrectly — which KBOs might have formed and evolved *in situ*.

Classical KBOs have spectral properties distinct from those of other dynamical classes: Their colors are more uniformly red (Fig. 3) (see chapter by Cruikshank et al.). According to the Kolmogorov-Smirnov test (*Press et al.,* 1992; *Peixinho et al.,* 2004), the probabilities that classical KBOs have B-V colors and Boehnhardt-S slopes (*Boehnhardt et al.,* 2001) drawn from distributions identical to those for resonant KBOs are 10^{-2} and 10^{-3}, respectively. When classical KBOs are compared to scattered KBOs (see below), the corresponding probabilities are 10^{-6} and 10^{-4}. An alternative interpretation is that low-i KBOs are redder than high-i KBOs (*Trujillo and Brown,* 2002; *Peixinho et al.,* 2004). This last claim is statistically significant when classical, scattered, and resonant KBOs are combined and analyzed as one set (Fig. 4). However, no correlation between physical properties and i (or any other measure of excitation) has proven significant within any individual class.

Both the inner edge of the classical belt at $a \approx 37$ AU, and the gap in the classical belt at $a \approx 40$–42 AU and $\langle i \rangle \leq 10°$ (see Fig. 2), reflect ongoing sculpting by the present-day planets. The inner edge marks the distance out to which the planets have eroded the Kuiper belt over the last few billion years (*HW93*; *Duncan et al.,* 1995, hereafter *D95*). The gap is carved by the ν_{18}, ν_{17}, and ν_8 secular resonances (*HW93*; *D95*; *Knežević et al.,* 1991). At a secular resonance denoted by ν_j, the orbital precession frequency of a test particle — apsidal if $j < 10$ and nodal if $j > 10$ — matches one of the precession eigenfrequencies of the planets (see Chapter 6 of *Murray and Dermott,* 1999). For example, at

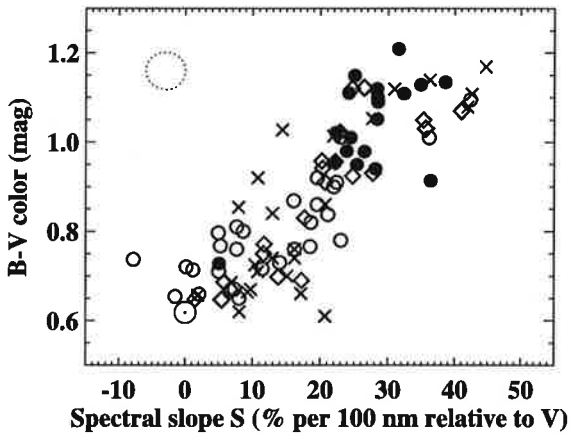

Fig. 3. Visual colors of KBOs and Centaurs calculated from published photometry, with the average uncertainty indicated by the upper left oval. The spectral slope S is calculated for wavelengths in the range of Johnson V through I. Neutral (solar) colors are indicated by ⊙. Symbols for dynamical classes are the same as those for Fig. 1. Classical KBOs constitute a distinct red cluster, except for (35671), which also has a small semimajor axis of 38 AU. Other classes are widely dispersed in color.

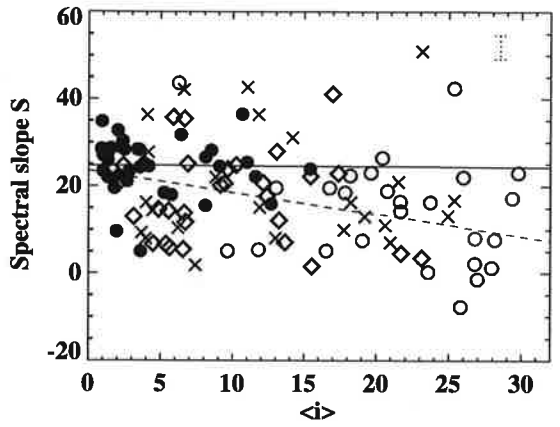

Fig. 4. Spectral slope S vs. time-averaged inclination. The typical uncertainty in S is indicated by the dotted bar. Classical KBOs evince no trend of color with ⟨i⟩. The solid line is fitted to classicals only; statistical tests using Spearman's rank-order coefficient and Kendall's tau (*Press et al.,* 1992) show that no significant correlation exists. When classical, resonant, and scattered KBOs are combined, S and ⟨i⟩ correlate significantly (with a false alarm probability of 10^{-5}); the dashed line is a fit to all three classes. The two most neutral classicals are (35671) and 1998 WV_{24}, having semimajor axes of 38 and 39 AU, respectively.

low i, the ν_8 resonance drives e to Neptune-crossing values in ~10^6 yr. Particles having large i, however, can elude the ν_8 (*Knežević et al.,* 1991). Indeed, 18 KBOs of various classes and all having large ⟨i⟩ reside within the gap.

By contrast, the outer edge of the classical belt at a ≈ 47 AU is likely primordial. Numerous surveys (e.g., *E05*; *Bernstein et al.,* 2004; and references therein) carried out after an edge was initially suspected (*Jewitt et al.,* 1998) all failed to find a single object moving on a low-e orbit outside 47 AU. The reality of the "Kuiper Cliff" is perhaps most convincingly demonstrated by *Trujillo and Brown* (2001), who simply plot the distribution of heliocentric discovery distances of (mostly classical) KBOs after correcting for the bias against finding more distant, fainter objects. This distribution peaks at 44 AU and plummets to a value 10× smaller at 50 AU. The statistical significance of the Cliff hinges upon the fact that the bias changes less dramatically — only by a factor of 2.2–2.4 for reasonable parameterizations of the size distribution — between 44 and 50 AU. The possibility remains that predominantly small objects having radii R < 50 km reside beyond 47 AU, or that the Cliff marks the inner edge of an annular gap having radial width ≥30 AU (*Trujillo and Brown,* 2001).

4. *Scattered KBOs* (106/529) comprise nonclassical, nonresonant objects whose perihelion distances q remain outside the orbit of Neptune. [The "scattered-near" and "scattered-extended" classes defined in *E05* — see also *Gladman et al.* (2002) — are combined to simplify discussion. Also, while we do not formally introduce Oort cloud objects as a class, we make connections to that population throughout this review.] How were scattered KBOs emplaced onto their highly elongated and inclined orbits? Appealing to perturbations exerted by the giant planets in their current orbital configuration is feasible only for some scattered objects. A rule of thumb derived from numerical experiments for the extent of the planets' collective reach is q ≤ 37 AU (*D95*; *Gladman et al.,* 2002). Figure 1 reveals that many scattered objects possess q > 37 AU and are therefore problematic. Outstanding examples include 2000 CR_{105} (q = 44 AU) (*Millis et al.,* 2002; *Gladman et al.,* 2002) and (90377) Sedna (q = 76 AU) (*Brown et al.,* 2004).

These classifications are intended to sharpen analyses and initiate discussion. The danger lies in allowing them to unduly color our thinking about origins. For example, although Sedna is classified above as a scattered KBO, the history of its orbit may be distinct from those of other scattered KBOs. We make this distinction explicit below.

2.2. Sky Density and Mass

We provide estimates for the masses of the Kuiper belt (comprising objects having q ≤ 60 AU and a > 30 AU; section 2.2.1); the inner Oort cloud (composed of Sedna-like objects; section 2.2.2); and Neptune Trojans (a ≈ 30 AU; section 2.2.3).

2.2.1. Main Kuiper belt. *Bernstein et al.* (2004, hereafter *B04*) compile data from published surveys in addition to their own unprecedentedly deep Hubble Space Telescope (HST) survey to compute the cumulative sky density of KBOs vs. apparent red magnitude m_R ("luminosity function"), shown in Fig. 5. Sky densities are evaluated near

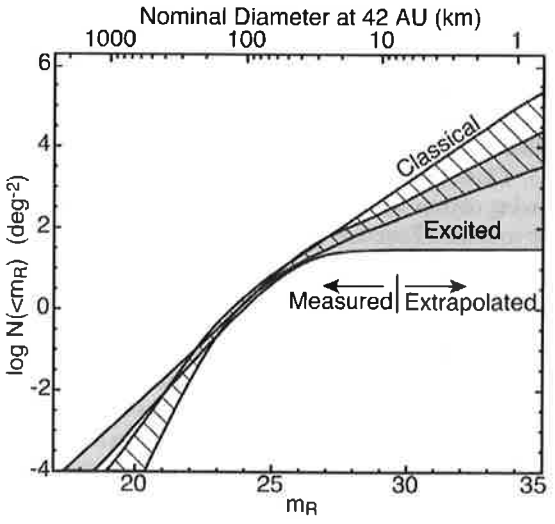

Fig. 5. Cumulative sky density vs. apparent red magnitude for CKBOs and excited KBOs, from *B04*. If $N(<m_R) \propto 10^{\alpha m_R}$, then the size distribution index $\tilde{q} = 5\alpha + 1$ (see section 2.2.1). Envelopes enclose 95% confidence intervals. The top abscissa is modified from *B04*; here it assumes a visual albedo of 10%. Figure courtesy of G. Bernstein.

the ecliptic plane. Objects are divided into two groups: "CKBOs" (similar to our classical population) having heliocentric distances 38 AU < d < 55 AU and ecliptic inclinations $i \leq 5°$, and "excited" KBOs (similar to our combined resonant and scattered classes) having 25 AU < d < 60 AU and $i > 5°$. Given these definitions, their analysis excludes objects with ultrahigh perihelia such as Sedna. With 96% confidence, *B04* determine that CKBOs and excited KBOs have different luminosity functions. Moreover, neither function conforms to a single power law from $m_R = 18$ to 29; instead, each is well-fitted by a double power law that flattens toward fainter magnitudes. The flattening occurs near $m_R \approx 24$ for both groups. To the extent that all objects in a group have the same albedo and are currently located at the same heliocentric distance, the luminosity function is equivalent to the size distribution. We define $\tilde{q}$ as the slope of the differential size distribution, where $dN \propto R^{-\tilde{q}}dR$ equals the number of objects having radii between R and R + dR. As judged from Fig. 5, for CKBOs, $\tilde{q}$ flattens from 5.5–7.7 (95% confidence interval) to 1.8–2.8 as R decreases. For excited KBOs, $\tilde{q}$ flattens from 4.0–4.6 to 1.0–3.1. Most large objects are excited (see also Fig. 1).

By integrating the luminosity function over all magnitudes, *B04* estimate the total mass in CKBOs to be

$$M_{CKBO} \approx 0.005 \left(\frac{p}{0.10}\right)^{-3/2} \left(\frac{d}{42\ AU}\right)^{6} \times \left(\frac{\rho}{2\ g\ cm^{-3}}\right)\left(\frac{A}{360° \times 6°}\right) M_\oplus \quad (2)$$

where all CKBOs are assumed to have the same albedo p, heliocentric distance d, and internal density ρ. The solid angle subtended by the belt of CKBOs is A. Given uncertainties in the scaling variables — principally p and ρ (see chapter by Cruikshank et al. for recent estimates) — this mass is good to within a factor of several. The mass is concentrated in objects having radii R ~ 50 km, near the break in the luminosity function.

The mass in excited KBOs cannot be as reliably calculated. This is because the sample is heterogeneous — comprising both scattered and resonant KBOs having a wide dispersion in d — and because corrections need to be made for the observational bias against finding objects near the aphelia of their eccentric orbits. The latter bias can be crudely quantified as $(Q/q)^{3/2}$: the ratio of time an object spends near its aphelion distance Q (where it is undetectable) vs. its perihelion distance q (where it is usually discovered). An order-of-magnitude estimate that accounts for how much larger A, d^6, and $(Q/q)^{3/2}$ are for excited KBOs than for CKBOs suggests that the former population might weigh ~10× as much as the latter, or ~0.05 $M_\oplus$. This estimate assumes that $\tilde{q}$ for excited KBOs is such that most of the mass is concentrated near $m_R \approx 24$, as is the case for CKBOs. If $\tilde{q}$ for the largest excited KBOs is as small as 4, then our mass estimate increases by a logarithm to ~0.15 $M_\oplus$.

2.2.2. Inner Oort cloud (Sedna-like objects). What about objects with unusually high perihelia such as Sedna, whose mass is $M_S \sim 6 \times 10^{-4}(R/750\ km)^3\ M_\oplus$? The Caltech Palomar Survey searched f ~ 1/5 of the celestial sphere to discover one such object (*Brown et al.,* 2004). By assuming Sedna-like objects are distributed isotropically (not a forgone conclusion; see section 3.7), we derive an upper limit to their total mass of $M_S(Q/q)^{3/2}f^{-1} \sim 0.1\ M_\oplus$. If all objects on Sedna-like orbits obey a size spectrum resembling that of excited KBOs (*B04*), then we revise the upper limit to ~0.3 $M_\oplus$. The latter value is 20× smaller than the estimate by *Brown et al.* (2004); the difference arises from our use of a more realistic size distribution.

2.2.3. Neptune Trojans. The first Neptune Trojan, 2001 QR_{322} (hereafter "QR"), was discovered by the DES (*C03*). The distribution of DES search fields on the sky, coupled with theoretical maps of the sky density of Neptune Trojans (*Nesvorný and Dones,* 2002), indicate that N ~ 10–30 objects resembling QR librate on tadpole orbits about Neptune's forward Lagrange (L4) point (*C03*). Presumably a similar population exists at L5. An assumed albedo of 12–4% yields a radius for QR of R ~ 65–115 km. Spreading the inferred population of QR-like objects over the area swept out by tadpole orbits gives a surface mass density in a single Neptune Trojan cloud that approaches that of the main Kuiper belt to within factors of a few (*Chiang and Lithwick,* 2005). Large Neptune Trojans are at least comparable in number to large Jupiter Trojans and may outnumber them by a factor of ~10 (*C03*).

2.3. Binarity

Veillet et al. (2002) optically resolved the first binary (1998 WW_{31}) among KBOs having sizes R ~ 100 km. Over 20 binaries having components in this size range are now resolved. Components typically have comparable brightnesses and are separated (in projection) by 300–10^5 km (e.g., *Stephens and Noll,* 2006). These properties reflect observational biases against resolving binaries that are separated by ≤0.1" and that contain faint secondaries.

Despite these selection biases, *Stephens and Noll* (2006) resolved as many as 9 out of 81 KBOs (~10%) with HST. They further report that the incidence of binarity appears 4× higher in the classical disk than in other dynamical populations. It is surprising that so many binaries exist with components widely separated and comparably sized. A typical binary in the asteroid belt, by contrast, comprises dissimilar masses separated by distances only slightly larger than the primary's radius. Another peculiarity of KBO binaries is that components orbit one other with eccentricities on the order of unity. In addition, binary orbits are inclined relative to their heliocentric orbits at seemingly random angles. See *Noll* (2003) for more quantitative details.

Although close binaries cannot be resolved, their components can eclipse each other. *Sheppard and Jewitt* (2004) highlight a system whose lightcurve varies with large amplitude and little variation in color, suggesting that it is a near-

contact binary. They infer that at least 10% of KBOs are members of similarly close binaries.

Among the four largest KBOs having R ≈ 1000 km — 2003 UB_{313}, Pluto, 2005 FY_9, and 2003 EL_{61} — three are known to possess satellites. Secondaries for 2003 EL_{61} and 2003 UB_{313} are 5% and 2% as bright as their primaries, and separated by 49,500 and 36,000 km, respectively (*Brown et al.,* 2005b). In addition to harboring Charon (*Christy and Harrington,* 1978), Pluto possesses two, more distant companions having R ~ 50 km (*Weaver et al.,* 2006). The three satellites' orbits are nearly co-planar; their semimajor axes are about 19,600, 48,700, and 64,800 km; and their eccentricities are less than 1% (*Buie et al.,* 2006).

3. THEORETICAL TIMELINE

We now recount a possible history of transneptunian space. Throughout our narration, it is helpful to remember that the timescale for an object of size R and mass M to collide into its own mass in smaller objects is

$$t_{col} \sim \frac{M}{\dot{M}} \sim \frac{\rho R^3}{(\sigma/h)R^2 v_{rel}} \sim \frac{\rho R}{\sigma\Omega} \qquad (3)$$

where $\dot{M}$ is the rate at which mass from the surrounding disk impacts the object, σ is the disk's surface mass density (mass per unit face-on area) in smaller objects, v_{rel} is the relative collision velocity, $h \sim v_{rel}/\Omega$ is the effective vertical scale height occupied by colliders, and Ω is the orbital angular frequency. Relative velocities v_{rel} depend on how e and i are distributed. Equation (3) requires that e's and i's be comparably distributed and large enough that gravitational focusing is ignorable. While these conditions are largely met by currently observed KBOs, they were not during the primordial era. Our expressions below represent appropriate modifications of equation (3).

3.1. Coagulation

The mass inferred for the present-day Kuiper belt, ~0.05–0.3 $M_\oplus$ (section 2.2), is well below that thought to have been present while KBOs coagulated. *Kenyon and Luu* (1998, 1999) and *Kenyon* (2002), in a series of particle-in-a-box accretion simulations, find that ~3–30 $M_\oplus$ of primordial solids, spread over an annulus extending from 32 to 38 AU, are required to (1) coagulate at least one object as large as Pluto and (2) coagulate ~10^5 objects having R > 50 km. The required initial surface density, $\sigma \sim 0.06$–0.6 g cm^{-2}, is on the order of that of the condensible portion of the minimum-mass solar nebula (MMSN) at 35 AU: $\sigma_{MMSN} \sim 0.2$ g cm^{-2}.

3.1.1. The missing-mass problem. That primordial and present-day masses differ by 2 orders of magnitude is referred to as the "missing-mass" problem. The same accretion simulations point to a possible resolution: Only ~1–2% of the primordial mass accretes to sizes exceeding ~100 km. The remainder stalls at comet-like sizes of ~0.1–10 km. Stunting of accretion is attributed to the formation of several Pluto-sized objects whose gravity amplifies velocity dispersions so much that collisions between planetesimals are erosive rather than accretionary. Thus, accretion in the Kuiper belt may be self-limiting (*Kenyon and Luu,* 1999). The bulk of the primordial mass, stalled at cometary sizes, is assumed by these authors to erode away by destructive collisions over gigayear timescales.

We can verify analytically some of Kenyon and Luu's results by exercising the "two-groups method" (*G04*), whereby the spectrum of planetesimal masses is approximated as bimodal. "Big" bodies each of size R, mass M, Hill radius R_H, and surface escape velocity v_{esc} comprise a disk of surface density Σ. They are held primarily responsible for stirring and accreting "small" bodies of size s, surface density σ, and random velocity dispersion u. By random velocity we mean the noncircular or nonplanar component of the orbital velocity. As such, u is proportional to the root-mean-squared dispersion in e and i.

To grow a big body takes time

$$t_{acc} \equiv \frac{R}{\dot{R}} \sim \frac{\rho R}{\sigma\Omega}\left(\frac{u}{v_{esc}}\right)^2 \qquad (4)$$

where the term in parentheses is the usual gravitational focussing factor (assumed <1). Gravitational stirring of small bodies by big ones balances damping of relative velocities by inelastic collisions among small bodies. This balance sets the equilibrium velocity dispersion u

$$\frac{\rho R}{\Sigma\Omega}\left(\frac{u}{v_{esc}}\right)^4 \sim \frac{\rho s}{\sigma\Omega} \qquad (5)$$

Combining equations (4) and (5) implies

$$t_{acc} \sim 10\left(\frac{R}{100\text{ km}}\frac{s}{100\text{ m}}\frac{\Sigma/\sigma}{0.01}\right)^{1/2}\left(\frac{\sigma_{MMSN}}{\sigma}\right)\text{ m.y.} \qquad (6)$$

$$u \sim 6\left(\frac{s}{100\text{ m}}\frac{\Sigma/\sigma}{0.01}\right)^{1/4}\left(\frac{R}{100\text{ km}}\right)^{3/4}\text{ m s}^{-1} \qquad (7)$$

at a distance of 35 AU. All bodies reside in a remarkably thin, dynamically cold disk: Eccentricities and inclinations are at most on the order of $u/(\Omega a) \sim 0.001$. Our nominal choices for σ, Σ, and s are informed by Kenyon and Luu's proposed solution to the missing-mass problem. Had we chosen values resembling those of the Kuiper belt today — $\Sigma \sim \sigma \sim 0.01\sigma_{MMSN}$ — coagulation times would exceed the age of the solar system.

The above framework for understanding the missing-mass problem, while promising, requires development. First,

account needs to be made for how the formation of Neptune — and possibly other planet-sized bodies — influences the coagulation of KBOs. None of the simulations cited above succeeds in producing Neptune-mass objects. Yet minimum-mass disks may be capable, in theory if not yet in simulation, of producing several planets having masses approaching that of Neptune at distances of 15–25 AU on timescales much shorter than the age of the solar system (*G04*) (see sections 3.4–3.5). The inability of simulations to produce ice giants may arise from their neglect of small-s, low-u particles that can be efficiently accreted (*G04*). Sizes s as small as centimeters seem possible. How would their inclusion, and the consequent formation of Neptune-mass planets near the Kuiper belt, change our understanding of the missing-mass problem? Second, how does the outer solar system shed ~99% of its primordial solids? The missing-mass problem translates to a "clean-up" problem, the solution to which will involve some as yet unknown combination of collisional comminution, diffusive transport by interparticle collisions, gravitational ejection by planets, and removal by gas and/or radiation drag.

3.1.2. The outer edge of the primordial planetesimal disk. How far from the Sun did planetesimals coagulate? The outer edge of the classical disk at 47 AU (section 2) suggests that planetesimals failed to form outside this distance. Extrasolar disks are also observed to have well-defined boundaries. The debris disks encircling β Pictoris and AU Microscopii exhibit distinct changes ("breaks") in the slopes of their surface brightness profiles at stellocentric distances of 100 AU and 43 AU, respectively (e.g., *Kalas and Jewitt,* 1995; *Krist et al.,* 2005). This behavior can be explained by having dust-producing parent bodies reside only at distances interior to break radii (*Strubbe and Chiang,* 2006).

We cannot predict with confidence how planetesimal disks truncate. Our understanding of how micrometer-sized dust assembles into the "small," super-meter-sized bodies that coagulation calculations presume as input is too poor. Recent work discusses how solid particles might drain toward their central stars by gas drag, and how the accumulation of such solids at small stellocentric distances triggers self-gravitational collapse and the formation of larger bodies (*Youdin and Shu,* 2002; *Youdin and Chiang,* 2004). These ideas promise to explain why planetesimal disks have sharp outer edges, but are subject to uncertainties regarding the viability of gravitational instability in a turbulent gas. To sample progress on planetesimal formation, see *Garaud and Lin* (2004), *Youdin and Goodman* (2005), and *Gómez and Ostriker* (2005). In what follows, we assume that objects having R ~ 100 km coagulated only inside 47 AU.

3.2. Formation of Binaries

To have formed from a fragmentary collision, binary components observed today cannot have too much angular momentum. Consider two big bodies undergoing a gravitationally focused collision. Each body has radius R, mass M, and surface escape velocity v_{esc}. Prior to the collision, their angular momentum is at most $L_{max} \sim Mv_{esc}R$. After the collision, the resultant binary must have angular momentum $L < L_{max}$. Unless significant mass is lost from the collision, components can be comparably sized only if their separation is comparable to their radii. Pluto and Charon meet this constraint. Their mass ratio is ~1/10, their separation is ~20 R_{Pluto}, and hence their angular momentum obeys $L/L_{max} \sim \sqrt{20}/10 \leq 1$. *Canup* (2005) explains how Charon might have formed by a collision. The remaining satellites of Pluto (*Stern et al.,* 2006), the satellites of Pluto-sized KBOs 2003 EL_{61} and 2003 UB_{313}, and the candidate near-contact binaries discovered by *Sheppard and Jewitt* (2004) might also have formed by collisions.

By contrast, binary components having wide separations and comparable masses have too much angular momentum to have formed by gravitationally focused collisions. And if collisions were unfocused, collision times would exceed the age of the solar system — assuming, as we do throughout this review, that the surface density of big (R ~ 100 km) bodies was the same then as now (section 1; section 3.9).

Big bodies can instead become bound ("fuse") by purely gravitational means while they are still dynamically cold. Indeed, such binaries testify powerfully to the cold state of the primordial disk. To derive our expressions below, recall that binary components separated by the Hill radius ~R_H orbit each other with the same period that the binary's center of mass orbits the Sun, ~Ω^{-1}. Furthermore, we assume that the velocity dispersion v of big bodies is less than their Hill velocity $v_H \equiv \Omega R_H$. Then big bodies undergo runaway cooling by dynamical friction with small bodies and settle into an effectively two-dimensional disk (*G04*). Reaction rates between big bodies must be calculated in a two-dimensional geometry. Because $u > v_H$, reaction rates involving small bodies take their usual forms appropriate for three dimensions.

Goldreich et al. (2002, hereafter *G02*) describe two collisionless formation scenarios, dubbed L^3 and L^2s. Both begin when one big body (L) enters a second big body's (L) Hill sphere. Per big body, the entry rate is

$$\dot{N}_H \sim \frac{\Sigma\Omega}{\rho R}\left(\frac{R_H}{R}\right)^2 \sim \frac{\Sigma\Omega}{\rho R}\alpha^{-2} \qquad (8)$$

where $\alpha \equiv R/R_H \approx 1.5 \times 10^{-4}(35\ \text{AU/a})$. If no other body participates in the interaction, the two big bodies would pass through their Hill spheres in a time Ω^{-1} (assuming they do not collide). The two bodies fuse if they transfer enough energy to other participants during the encounter. In L^3, transfer is to a third big body: $L + L + L \rightarrow L^2 + L$. To just bind the original pair, the third body must come within R_H of the pair. The probability for this to happen in time Ω^{-1} is $P_{L^3} \sim \dot{N}_H\Omega^{-1}$. If the third body succeeds in approaching this close, the probability that two bodies fuse is on the or-

der of unity. Therefore the timescale for a given big body to fuse to another by L^3 is

$$t_{fuse,L^3} \sim \frac{1}{\dot{N}_H P_{L^3}} \sim \left(\frac{\rho R}{\Sigma}\right)^2 \frac{\alpha^4}{\Omega} \sim 2 \text{ m.y.} \quad (9)$$

where the numerical estimate assumes R = 100 km, $\Sigma = 0.01\sigma_{MMSN}$, and a = 35 AU.

In L^2s, energy transfer is to small bodies by dynamical friction: $L + L + s^\infty \rightarrow L^2 + s^\infty$. In time $\sim\Omega^{-1}$, the pair of big bodies undergoing an encounter lose a fraction $(\sigma\Omega/\rho R)(v_{esc}/u)^4\Omega^{-1}$ of their energy, under the assumption $v_{esc} > u > v_H$ (*G04*). This fraction is on the order of the probability P_{L^2s} that they fuse, whence

$$t_{fuse,L^2s} \sim \frac{1}{\dot{N}_H P_{L^2s}} \sim \left(\frac{\rho R}{\sigma}\right)^2 \frac{s}{R}\frac{\alpha^2}{\Omega} \sim 7 \text{ m.y.} \quad (10)$$

where we have used equation (5) and set s = 100 m.

Having formed with semimajor axis $x \sim R_H \sim 7000$ R, a binary's orbit shrinks by further energy transfer. If L^3 is the more efficient formation process, passing big bodies predominantly harden the binary; if L^2s is more efficient, dynamical friction dominates hardening. The probability P per orbit that x shrinks from $\sim R_H$ to $\sim R_H/2$ is on the order of either P_{L^3} or P_{L^2s}. We equate the formation rate of binaries, N_{all}/t_{fuse}, with the shrinkage rate, $\Omega P N_{bin}|_{x \sim R_H}$, to conclude that the steady-state fraction of KBOs that are binaries with separation R_H is

$$f_{bin}(x \sim R_H) \equiv \left.\frac{N_{bin}}{N_{all}}\right|_{x \sim R_H} \sim \frac{\Sigma}{\rho R}\alpha^{-2} \sim 0.4\% \quad (11)$$

As x decreases, shrinkage slows. Therefore f_{bin} increases with decreasing x. Scaling relations can be derived by arguments similar to those above. If L^2s dominates, $f_{bin} \propto x^0$ for $x > R_H(v_H/u)^2$ and $f_{bin} \propto x^{-1}$ for $x < R_H(v_H/u)^2$ (*G02*). If L^3 dominates, $f_{bin} \propto x^{-1/2}$.

Alternative formation scenarios require, in addition to gravitational scatterings, physical collisions. *Weidenschilling* (2002) suggests a variant of L^3 in which the third big body collides with one member of the scattering pair. Since physical collisions have smaller cross-sections than gravitational interactions, this mechanism requires $\sim 10^2$ more big (R ~ 100 km) bodies than are currently observed to produce the same rate that is cited above for L^3 (*Weidenschilling,* 2002). *Funato et al.* (2004) propose that observed binaries form by the exchange reaction $Ls + L \rightarrow L^2 + s$: A small body of mass m, originally orbiting a big body of mass M, is ejected by a second big body. In the majority of ejections, the small body's energy increases by its orbital binding energy $\sim mv_{esc}^2/2$, leaving the big bodies bound to each other with separation $x \sim (M/m)R$. The rate-limiting step is the formation of the preexisting (Ls) binary, which requires (as in the asteroid belt) two big bodies to collide and fragment. Hence

$$t_{fuse,exchange} \sim \frac{\rho R}{\Sigma\Omega}\alpha^{3/2} \sim 0.6 \text{ m.y.} \quad (12)$$

Estimating f_{bin} as a function of x under the exchange hypothesis requires knowing the distribution of fragment masses m. Whether L^2s, L^3, or exchange reactions dominate depends on the uncertain parameters Σ, σ, and s.

As depicted above, newly formed binaries should be nearly co-planar with the two-dimensional disk of big bodies, i.e., binary orbit normals should be nearly parallel. Observations contradict this picture (section 2.3). How dynamical stirring of the Kuiper belt subsequent to binary formation affects binary inclinations and eccentricities has not been investigated.

3.3. Early Stirring by Growing Planetary Oligarchs

Coagulation of KBOs and fusing of binaries cannot proceed today, in part because velocity dispersions are now so large that gravitational focusing is defeated on a wide range of length scales. What stirred the Kuiper belt? There is no shortage of proposed answers. Much of the remaining review (sections 3.3–3.7) explores the multitude of nonexclusive possibilities. We focus on stirring "large" KBOs like those currently observed, having R ~ 100 km. Our setting remains the primordial disk, of whose mass large KBOs constitute only a small fraction (1–2%; section 3.1.1).

Neptune and Uranus are thought to accrete as oligarchs, each dominating their own annulus of full-width ~5 Hill radii (*G04*; *Ida and Makino,* 1993; *Greenberg et al.,* 1991). (The coefficient of 5 presumes that oligarchs feed in a shear-dominated disk in which planetesimals have random velocities u that are less than the oligarch's Hill velocity $v_H = \Omega R_H$. If $u > v_H$, oligarchs' feeding annuli are wider by $\sim u/v_H$. In practice, u/v_H does not greatly exceed unity since it scales weakly with input parameters.) Each oligarch grows until its mass equals the isolation mass

$$M_p \sim 2\pi a \times 5R_{H,p} \times \sigma \quad (13)$$

where $R_{H,p}$ is the oligarch's Hill radius. For a = 25 AU and M_p equal to Neptune's mass $M_N = 17\ M_\oplus$, equation (13) implies $\sigma \sim 0.9$ g cm$^{-2} \sim 3\sigma_{MMSN}$. About 5 Neptune-mass oligarchs can form in nested annuli between 15 and 25 AU. Inspired by *G04* who point out the ease with which ice giants coagulate when the bulk of the disk mass comprises very small particles (section 3.1.1), we assume that all five do form in a disk that is a few times more massive than the MMSN and explore the consequences of such an initially packed system.

While oligarchs grow, they stir large KBOs in their immediate vicinity. A KBO that comes within distance b of mass M_p has its random velocity excited to $v_K \sim (GM_p/b)^{1/2}$. Take the surface density of perturbers to be Σ_p. Over time t,

a KBO comes within distance $b \sim [M_p/(\Sigma_p \Omega t)]^{1/2}$ of a perturber. Therefore

$$v_K \sim G^{1/2}(M_p \Sigma_p \Omega t)^{1/4} \quad (14)$$

Since Neptune and Uranus contain more hydrogen than can be explained by accretion of icy solids alone, they must complete their growth within $t_{acc,p} \sim$ 1–10 m.y., before all hydrogen gas in the MMSN photoevaporates (e.g., *Matsuyama et al.,* 2003, and references therein). For $t = t_{acc,p} =$ 10 m.y., $M_p = M_N$, $\Sigma_p = 0.9$ g cm^{-2}, and $\Omega = 2\pi/(100$ yr), equation (14) implies $v_K \sim 1$ km s^{-1} or $e_K \sim 0.2$. It is safe to neglect damping of v_K for large KBOs, which occurs by inelastic collisions over a timescale $t_{col} \sim 400$ (0.9 g cm^{-2}/ σ) m.y. $\gg t_{acc,p}$.

3.4. Velocity Instability and Ejection of Planets

Once the cohort of Neptune-mass oligarchs consumes approximately one-half the mass of the parent disk, they scatter one another onto highly elliptical and inclined orbits (equation (111) of *G04*; *Kenyon and Bromley,* 2006). This velocity instability occurs because damping of planetary random velocities by dynamical friction with the disk can no longer compete with excitation by neighboring, crowded oligarchs.

The epoch of large planetary eccentricities lasts until enough oligarchs are ejected from the system. We can estimate the ejection time by following the same reasoning that led to equation (14). Replace v_K with the system escape velocity $v_{esc,sys} \sim \Omega a$, and replace Σ_p with the surface density of oligarchs $\sim M_p/a^2$ (see equation (13)). Then solve for

$$t = t_{eject} \sim \left(\frac{M_\odot}{M_p}\right)^2 \frac{0.1}{\Omega} \quad (15)$$

The coefficient of 0.1 is attributed to more careful accounting of encounter geometries; equation (15) gives ejection times similar to those found in numerical simulations (*G04*). Neptune-mass oligarchs at $a \approx 20$ AU kick their excess brethren out over $t_{eject} \sim 600$ m.y. Removal is faster if excess oligarchs are passed inward to Jupiter and Saturn.

Oligarchs moving on eccentric orbits likely traverse distances beyond 30 AU and stir KBOs. We expect more members are added to the scattered KBO disk during this stage.

We have painted a picture of dynamically hot oligarchs similar to that drawn by *Thommes et al.* (1999) (see also *Tsiganis et al.,* 2005), who hypothesize that Neptune and Uranus form as oligarchs situated between the cores of Jupiter and Saturn at 5–10 AU. The nascent ice giants are scattered outward onto eccentric orbits once the gas giant cores amass their envelopes. While Neptune and Uranus reside on eccentric orbits, they can stir KBOs in much the same way as we have described above (*Thommes et al.,* 2002). Despite the similarity of implications for the stirring of KBOs, the underlying motivation of the cosmogony proposed by *Thommes et al.* (1999) is the belief that Neptune-mass bodies do not form readily at distances of ~30 AU. Recent work highlighting the importance of inelastic collisions among very small bodies challenges this belief (*G04*) (see section 3.1).

3.5. Dynamical Friction Cooling of Surviving Planets

Planetary oligarchs that survive ejection — i.e., Uranus and Neptune — have their e's and i's restored to small values by dynamical friction with the remnant disk (comprising predominantly small KBOs of surface density σ and velocity dispersion u) over time

$$t_{df,cool} = \frac{v_p}{\dot{v}_p} \sim \frac{\rho R_p}{\sigma \Omega}\left(\frac{v_p}{v_{esc,p}}\right)^4 \quad (16)$$

where R_p, $v_{esc,p}$, and $v_p \gg u$ are the planet's radius, surface escape velocity, and random velocity, respectively. For $v_p = \Omega a/2$ (planetary eccentricity $e_p \sim 0.5$), $a = 25$ AU, $R_p =$ 25000 km, $v_{esc,p} = 24$ km s^{-1}, and $\sigma = \Sigma_p = 0.9$ g cm^{-2} (since the velocity instability occurs when the surface density of oligarchs equals that of the parent disk; section 3.4), we find $t_{df,cool} \sim 20$ m.y.

While Neptune's orbit is eccentric, the planet might repeatedly invade the Kuiper belt at $a \approx$ 40–45 AU and stir KBOs. Neptune would have its orbit circularized by transferring energy to both small and large KBOs. Unlike small KBOs, large ones cannot shed this energy because they cool too inefficiently by inelastic collisions (see the end of section 3.3). Insert equation (16) into equation (14) and set $\Sigma_p = \sigma$ to estimate the random velocity to which large KBOs are excited by a cooling Neptune

$$v_K \sim v_p \quad (17)$$

Thus large KBOs are stirred to the same random velocity that Neptune had when the latter began to cool, regardless of the numerical value of $t_{df,cool}$. Large KBOs effectively record the eccentricity of Neptune just prior to its cooling phase. Final eccentricities e_K might range from ~0.1 to nearly 1. During this phase, the population of the scattered KBO disk would increase, perhaps dramatically so. If all large KBOs are stirred to $e_K \gg 0.1$, new large KBOs must coagulate afterward from the remnant disk of small, dynamically cold bodies to reconstitute the cold classical disk. Cold classicals might therefore postdate hot KBOs.

3.6. Planetary Migration

Having seen a few of its siblings evicted, and having settled onto a near-circular, flattened orbit, Neptune remains immersed in a disk of small bodies. The total mass of the disk is still a few times that of the planet because the prior

velocity instability occurred when the surface density of oligarchs was comparable to that of the disk. By continuing to scatter small bodies, Neptune migrates: Its semimajor axis changes while its eccentricity is kept small by dynamical friction. Absent other planets, migration would be sunward on average. Planetesimals repeatedly scattered by Neptune would exchange angular momentum with the planet in a random-walk fashion. Upon gaining specific angular momentum $\sim(\sqrt{2}-1)\Omega a^2$, where Ω and a are appropriate to Neptune's orbit, a planetesimal initially near Neptune would finally escape. Having lost angular momentum to the ejected planetesimal, Neptune would migrate inward. [A single planet can still migrate outward if it scatters material having predominantly higher specific angular momentum. *Gomes et al.* (2004) achieve this situation by embedding Neptune in a disk whose mass is at least 100 $M_\oplus$ and is weighted toward large distances ($\sigma a^2 \propto a$); see also the chapter by Morbidelli et al.).]

Other planets complicate this process. Numerical simulations by *Fernández and Ip* (1984) and *Hahn and Malhotra* (1999) incorporating all four giant planets reveal that planetesimals that originate near Neptune are more likely ejected by Jupiter. Over the course of their random walks, planetesimals lose angular momentum to Neptune and thereby cross Jupiter's orbit. Jupiter summarily ejects them (see equation (15) and related discussion). Thus, on average, Neptune gains angular momentum and migrates outward, as do Saturn and Uranus, while Jupiter's orbit shrinks.

An outward-bound Neptune passes objects to the interior planets for eventual ejection and seeding of the Oort cloud. We refer to this process as "scouring" the transneptunian disk. Scouring and migration go hand in hand; the fraction by which Neptune's semimajor axis increases is on the order of the fraction that the disk mass is scoured. Scouring is likely a key part of the solution to the clean-up (a.k.a. missing-mass) problem. If clean-up is not achieved by the end of Neptune's migration, one must explain how to transport the bulk of the transneptunian disk to other locales while keeping Neptune in place (*Gomes et al.*, 2004). Scouring has only been treated in collisionless N-body simulations. How scouring and migration proceed in a highly collisional disk of small bodies is unknown (section 3.6.4). In addition to scouring the disk, Neptune's migration has been proposed to sculpt the disk in other ways — by capturing bodies into mean-motion resonances (section 3.6.1), redistributing the classical disk by resonance capture and release (section 3.6.2), and deflecting objects onto scattered orbits (section 3.6.3). We critically examine these proposals below.

3.6.1. Capture and excitation of resonant Kuiper belt objects. As Neptune migrates outward, its exterior mean-motion resonances (MMRs) sweep across transneptunian space. Provided the migration is sufficiently slow and smooth, MMRs may trap KBOs and amplify their orbital eccentricities and, to a lesser extent, their inclinations. The eccentric orbits of Pluto and the Plutinos — objects that all inhabit Neptune's 3:2 resonance — may have resulted from resonance capture and excitation by a migrating Neptune (*Malhotra,* 1993, 1995; *Jewitt and Luu,* 2000). The observed occupation of other low-order resonances — e.g., the 4:3, 5:3, and 2:1 MMRs — by KBOs on eccentric orbits (see Fig. 2 and Table 1) further support the migration hypothesis (*C03*). In this section, we review the basic mechanism of resonant excitation of eccentricity, examine how the migration hypothesis must change in light of the unexpected occupation of high-order (e.g., the 7:4, 5:2, and 3:1) MMRs, and discuss how m:1 resonances serve as speedometers for Neptune's migration.

Consider the interaction between a test particle (KBO) and a planet on an expanding circular orbit. In a frame of reference centered on the Sun and rotating with the planet's angular velocity $\Omega_p(t)$, the particle's Hamiltonian is

$$\mathcal{H} = \mathcal{E} - \Omega_p(t)\mathcal{L} - \mathcal{R}(t) \tag{18}$$

where $\mathcal{E} = -GM_\odot/2a$, $\mathcal{L} = [GM_\odot a(1-e^2)]^{1/2}$, and $\mathcal{R}$ is the disturbing potential due to the planet (these quantities should be expressed in canonical coordinates). From Hamiltonian mechanics, $d\mathcal{H}/dt = \partial\mathcal{H}/\partial t = -\dot{\Omega}_p\mathcal{L} - \partial\mathcal{R}/\partial t$. Therefore

$$\frac{d\mathcal{E}}{dt}(1-\varepsilon) - \Omega_p\frac{d\mathcal{L}}{dt} = 0 \tag{19}$$

where $\varepsilon \equiv (d\mathcal{R}/dt - \partial\mathcal{R}/\partial t)/(d\mathcal{E}/dt)$. We rewrite equation (19) as

$$\frac{de^2}{dt} = \frac{(1-e^2)^{1/2}}{a}[(1-e^2)^{1/2} - \Omega/\Omega_p(1-\varepsilon)]\frac{da}{dt} \tag{20}$$

where Ω is the particle's angular frequency.

For a particle trapped in m:n resonance (where m and n are positive, relatively prime integers), a, e, and the resonance angle change little over the particle's orbital period. If the synodic period is not much longer than the orbital period, we may average the Hamiltonian over the former (we may do this by choosing appropriate terms in the expansion of $\mathcal{R}$). This yields $\Omega/\Omega_p(1-\varepsilon) = n/m$. For a particle in resonance, $|\varepsilon| \ll 1$. By change of variable to $x \equiv (1-e^2)^{1/2}$, equation (20) integrates to

$$[(1-e^2)^{1/2} - n/m]^2\, a = \text{constant} \tag{21}$$

which relates changes in a to changes in e for any resonance — exterior $m > n$, interior $m < n$, or Trojan $m = n$. In the case of a planet that migrates toward a particle in exterior resonance, a increases to maintain resonant lock (*Goldreich,* 1965; *Peale,* 1986). Then by equation (21), e also tends to increase, toward a maximum value $[1-(n/m)^2]^{1/2}$. Particles inhabiting either an exterior or interior resonance have their eccentricities amplified from 0 because they are perturbed by a force pattern whose angular speed Ω_p does not equal their orbital angular speed Ω. Particles receive energy and angular momentum from the planet in a ratio that cannot maintain circularity of orbits.

Among observed 2:1 resonant KBOs, max(e) ≈ 0.38 (Fig. 2). If 2:1 resonant KBOs had their eccentricities amplified purely by migration, they must have migrated by $\Delta a \approx 13$ AU (equation (21)). Neptune must have migrated correspondingly by $\Delta a_p \approx 8$ AU. This is an upper bound on Δa_p because it does not account for nonzero initial eccentricities prior to capture.

In early simulations (*Malhotra,* 1993, 1995) of resonance capture by a migrating Neptune, resonances swept across KBOs having initially small e's and i's. These models predicted that if Neptune's orbit expanded by $\Delta a_p \approx 8$ AU, low-order resonances such as the 4:3, 3:2, 5:3, and 2:1 MMRs would be occupied by objects having $0.1 \lesssim e \lesssim 0.4$ and $i \lesssim 10°$. Eccentric KBOs indeed inhabit these resonances (Fig. 2). Two observations were not anticipated: (1) Resonant KBOs are inclined by up to $i \approx 30°$, and (2) high-order resonances — e.g., the 5:2, 7:4, and 3:1 — enjoy occupation. These observations suggest that Neptune's MMRs swept across not only initially dynamically cold objects, but also initially hot ones: The belt was preheated. For example, to capture KBOs into the 5:2 MMR, preheated eccentricities must be ≥0.1 (*C03*). Neptune-sized perturbers (sections 3.3–3.5) might have provided the requisite preheating in e and i.

To understand why capture into high-order resonances favors particles having larger initial e, recognize that capture is only possible if, over the time the planet takes to migrate across the maximum possible libration width $\max(\delta a_{lib})$, the particle completes at least one libration

$$\frac{\max(\delta a_{lib})}{|\dot{a}_p|} > T_{lib} \qquad (22)$$

where T_{lib} is the libration period (*Dermott et al.,* 1988). Otherwise, the particle would hardly feel the resonant perturbation as the planet races toward it. Since $\max(\delta a_{lib}) \sim (T_{orb}/T_{lib})a_p$ and $T_{lib} \sim T_{orb}(M_\odot e^{-|m-n|}/M_p)^{1/2}$ (*Murray and Dermott,* 1999), where T_{orb} is the orbital period of the particle and M_p is the mass of the planet, we rewrite equation (22) as

$$\frac{T_{lib}^2}{T_{orb}T_{mig}} \sim \frac{T_{orb}}{T_{mig}}\frac{M_\odot}{M_p}\frac{1}{e^{|m-n|}} < 1 \qquad (23)$$

where the migration timescale $T_{mig} \equiv a_p/|\dot{a}_p|$. The higher the order $|m - n|$ of the resonance, the greater e must be to satisfy equation (23) (*C03*; *Hahn and Malhotra,* 2005).

Asymmetric (m:1) resonances afford a way to estimate the migration timescale observationally. An asymmetric MMR furnishes multiple islands of libration. At the fixed point of each island, a particle's direct acceleration by Neptune balances its indirect acceleration by the Sun due to the Sun's reflex motion (*Pan and Sari,* 2004; *Murray-Clay and Chiang,* 2005, hereafter *MC05*). The multiplicity of islands translates into a multiplicity of orbital longitudes, measured relative to Neptune's, where resonant KBOs cluster on the sky. The pattern of clustering varies systematically with migration speed at the time of capture (*Chiang and Jordan,* 2002). For example, when migration is fast—occurring on timescales $T_{mig} \lesssim 20$ m.y. — objects are caught into 2:1 resonance such that more appear at longitudes trailing, rather than leading, Neptune's. The degree of asymmetry can be as large as 300%. When migration is slow, the distribution of captured 2:1 objects is symmetric about the Sun-Neptune line. The preference for trailing vs. leading longitudes arises from migration-induced shifts in the stable and unstable equilibria of the resonant potential. Shifts in the equilibrium values of the resonance angle are given in radians by equation (23) and are analogous to the shift in the equilibrium position of a spring in a gravitational field (*MC05*). The observation that trailing 2:1 KBOs do not outnumber leading ones constrains $T_{mig} > 20$ m.y. with nearly 3σ confidence (*MC05*). This measurement accords with numerical simulations of the migration process itself by *Hahn and Malhotra* (1999) and by *Gomes et al.* (2004, their Fig. 10); in these simulations, $T_{mig} \gtrsim 40$ m.y.

3.6.2. Stochastic migration and resonance retainment. Finite sizes of planetesimals render planetary migration stochastic ("noisy"). The numbers of high- and low-momentum objects that Neptune encounters over fixed time intervals fluctuate randomly. These fluctuations sporadically hasten and slow — and might occasionally even reverse — the planet's migration. Apportioning a fixed disk mass to larger (fewer) planetesimals generates more noise. Extreme noise defeats resonance capture. Therefore the existence of resonant KBOs — which we take to imply capture efficiencies on the order of unity — sets an upper limit on the sizes of planetesimals (small bodies) comprising the bulk of the mass of the disk. *Murray-Clay and Chiang* (2006, hereafter *MC06*) estimate this upper limit to be $s_{max} \sim O(100)$ km; a shortened derivation of their result reads as follows.

For a given planetesimal size, most noise is generated per unit mass disk by planetesimals having sub-Hill ($u < v_{H,p} = \Omega R_p/\alpha$) velocity dispersions and semimajor axes displaced $\pm R_{H,p}$ from the planet's (*MC06*). A single such planetesimal of mass μ, after undergoing a close encounter with the planet, changes the planet's semimajor axis by $\Delta a_1 \sim \pm(\mu/M_p)R_{H,p}$. The planet encounters such planetesimals at a rate $\dot{N} \sim \sigma R_{H,p}^2 \Omega_p/\mu$. Over the duration of migration $\sim(\Delta a_p/a_p)T_{mig}$, the planet's semimajor axis random walks away from its nominal (zero-noise) value by $\Delta a_{rnd} \sim \pm(\dot{N}\Delta a_p T_{mig}/a_p)^{1/2}|\Delta a_1|$. The libration amplitude in a of any resonant KBO increases by about this same $|\Delta a_{rnd}|$. Then stochasticity does not defeat resonance capture if $|\Delta a_{rnd}| < \max(\delta a_{lib})$; that is, if

$$s \lesssim \left(\frac{M_p}{M_\odot}\right)^{1/9}\left(\frac{\rho R_p e a_p}{\sigma \Omega_p T_{mig}\Delta a_p}\right)^{1/3}\alpha^{2/3}R_p \qquad (24)$$

which evaluates to $s \lesssim O(100)$ km for $a_p = 30$ AU, $\Delta a_p = 8$ AU, $T_{mig} \approx 40$ m.y., $\sigma = 0.2$ g cm^{-2}, and $e = 0.2$.

The above constraint on size applies to those planetesimals that comprise the bulk of the disk mass. Noise is also

introduced by especially large objects that constitute a small fraction of the disk mass. The latter source of noise has been invoked to explain the curious near-coincidence between the edge of the classical disk (a = 47 AU) and Neptune's 2:1 resonance (a = 47.8 AU). *Levison and Morbidelli* (2003) suggest that the sweeping 2:1 MMR captures KBOs only to release them en route because of close encounters between Neptune and objects having ~10× the mass of Pluto ("super-Plutos"). Dynamically cold KBOs, assumed to coagulate wholly inside 35 AU (section 3.1.2), are thereby combed outward to fill the space interior to the final location of the 2:1 MMR. Why the super-Plutos that are invoked to generate stochasticity have not been detected by wide-field surveys is unclear (*Morbidelli et al.,* 2002). The scenario further requires that ~3 $M_\oplus$ be trapped within the 2:1 MMR so that a secular resonance maintains a population of 2:1 resonant KBOs on low-e orbits during transport.

3.6.3. Contribution of migration to scattered Kuiper belt objects. Neptune migrates by scattering planetesimals. What fraction of these still reside today in the scattered belt? Do hot classicals (having $i \geq 5°$) owe their excitation to a migratory Neptune? Many scattered and hot classical KBOs observed today have q > 37 AU. This fact is difficult to explain by appealing to perturbers that reside entirely inside 30 AU. Insofar as a close encounter between a perturber and a particle can be modeled as a discontinous change in the particle's velocity at fixed position, the particle (assuming it remains bound to the Sun) tends to return to the same location at which it underwent the encounter.

Gomes (2003a,b) proposes that despite this difficulty, objects scattered by Neptune during its migration from ~20 to 30 AU can evolve into today's scattered and hot classical KBOs by having their perihelia raised by a variety of sweeping secular resonances (SRs; see section 2.1). As the outer planets migrate, SRs sweep across transneptunian space. After having its e and i amplified by close encounters with Neptune, a planetesimal may be swept over by an SR. Unlike MMRs, SRs cannot alter particle semimajor axes and therefore do not permanently trap particles. However, a particle that is swept over by an apsidal-type SR can have its eccentricity increased or decreased. A particle swept over by a secular resonance is analogous to an ideal spring of natural frequency ω_0, driven by a force whose time-variable frequency $\omega(t)$ sweeps past ω_0. Sweeping ω past ω_0 can increase or decrease the amplitude of the spring's free oscillation (the component of the spring's displacement that varies with frequency ω_0), depending on the relative phasing between driver and spring near the moment of resonance crossing when $\omega \approx \omega_0$.

Lowering e at fixed a raises q. *Gomes* (2003ab) and *Gomes et al.* (2005) find in numerical simulations of planetary migration that Neptune-scattered planetesimals originating on orbits inside 28 AU can have their perihelia raised up to 69 AU by a combination of sweeping SRs, MMRs, and Kozai-type resonances (which are a kind of SR). In addition to offering an explanation for the origin of high-q, high-i KBOs, this scenario also suggests a framework for understanding differences in physical properties between dynamical classes. Compared to classical KBOs, which are held to coagulate and evolve largely *in situ*, scattered KBOs originate from smaller heliocentric distances d. To the (unquantified) extents that coagulation rates and chemical environments vary from $d \approx 20$–50 AU, we can hope to understand why a large dispersion in i — which in the proposed scenario reflects a large dispersion in birth distance d — implies a large dispersion in color/size.

The main difficulty with this perihelion-raising mechanism is its low efficiency: Only ~0.1% of all objects that undergo close encounters with a migratory Neptune have their perihelia raised to avoid further close encounters over the age of the solar system (*Gomes,* 2003a,b). Based on this mechanism alone, a disk weighing ~50 $M_\oplus$ prior to migration would have ~0.05 $M_\oplus$ deposited into the scattered and hot classical belts for long-term storage. But only ~1–2% of this mass would be in bodies having sizes $R \geq 100$ km (*Kenyon and Luu,* 1998, 1999; *Kenyon,* 2002) (section 3.1).Therefore this scenario predicts that scattered and hot classical KBOs having $R \geq 100$ km would weigh, in total, ~10^{-3} $M_\oplus$ — about 50–150× below what is observed (section 2.2). This discrepancy is missed by analyses that neglect consideration of the KBO size distribution. A secondary concern is that current numerical simulations of this mechanism account for the gravitational effects of disk particles on planets but not on other disk particles. Proper calculation of the locations of secular resonances requires, however, a full accounting of the mass distribution.

Given the low efficiency of the mechanism, we submit that the high-q orbits of hot classical and scattered KBOs did not arise from Neptune's migration. Instead, these orbits may have been generated by Neptune-mass oligarchs whose trajectories passed through the Kuiper belt. While a numerical simulation is necessary to test this hypothesis, our order-of-magnitude estimates (sections 3.3–3.5) for the degree to which oligarchs stir the belt by simple close encounters are encouraging. No simulation has yet been performed in which the Kuiper belt is directly perturbed by a mass as large as Neptune's for a time as long as $t_{df,cool} \sim$ 20 m.y. Differences in physical properties between classical and scattered/resonant KBOs might still be explained along the same lines as described above: Scattered/resonant KBOs were displaced by large distances from their coagulation zones and so might be expected to exhibit a large dispersion in color and size, while classical KBOs were not so displaced. Even if all KBOs having $R \geq 100$ km were heated to large e or i by planetary oligarchs, the cold classical disk might have regenerated itself in a second wave of coagulation from a collisional disk of small bodies.

3.6.4. Problems regarding migration. The analyses of migration cited above share a common shortcoming: They assume that planetesimals are collisionless. But coagulation studies (section 3.1) indicate that much of the primordial mass remains locked in small bodies for which collision times threaten to be shorter than the duration of planetary migration. By equation (3), planetesimals having

sizes ≪1 km in a minimum-mass disk have collision times ≪20 m.y. How Neptune's migration unfolds when most of the disk comprises highly collisional bodies has not been well explored. Neptune may open a gap in the disk (in the same way that moons open gaps in collisional planetary rings) and the planet's migration may be tied to how the disk spreads by collisional diffusion (*Goldreich et al.,* 2004b).

How does the classical belt shed 99% of its primordial mass? Situated at 40–47 AU, it may be too distant for Neptune to scour directly. Perhaps the small bodies of the classical belt are first transported inward, either by gas drag or collisional diffusion, and subsequently scoured. Clean-up and migration are intertwined, but the processes are often not discussed together (but see *Gomes et al.,* 2004).

Are there alternatives to migration for the capture of resonant KBOs? Perhaps resonant KBOs are captured as Neptune's orbit cools by dynamical friction (section 3.5). Before capture, many belt members would already be stirred to large e and i, not only by unstable oligarchs (section 3.4), but also by Neptune while it cools. Cooling accelerates as it proceeds (equation (16)). A rapid change in the planet's semimajor axis toward the end of cooling might trap KBOs into resonance by serendipity. Just after Neptune's semimajor axis changes, objects having orbital elements (including longitudes) suitable for libration would be trapped. This speculative "freeze-in" mechanism might be too inefficient, since it requires that the fraction of phase-space volume occupied by resonances equal the fraction of KBOs that are resonant. Taken at face value, observations suggest the latter fraction is not much smaller than on the order of unity (section 2.1).

3.7. Stellar Encounters

A passing star may have emplaced Sedna onto its high-perihelion orbit. For the last t ~ 4 G.y., solar-mass stars in the solar neighborhood have had an average density $n_* \sim 0.04$ stars pc^{-3} and a velocity dispersion $\langle v_*^2 \rangle^{1/2} \sim 30$ km s^{-1}. If we assume that the Sun once resided within a "typical" open cluster, then $n_* \sim 4$ stars pc^{-3} and $\langle v_*^2 \rangle^{1/2} \sim 1$ km s^{-1} over t ~ 200 m.y. Over t ≥ 200 m.y., open clusters dissolve by encounters with molecular clouds (*Binney and Tremaine,* 1987). The number of stars that fly by the Sun within a distance q_* large enough that gravitational focussing is negligible ($q_* \geq GM_\odot / \langle v_*^2 \rangle \sim 900$ AU for $\langle v_*^2 \rangle^{1/2} \sim 1$ km s^{-1}) increases as $\int^t n_* \langle v_*^2 \rangle^{1/2} dt$. Therefore flybys during the current low-density era outnumber those during the cluster era by a factor of ~6. Nonetheless, intracluster encounters can be more effective at perturbing KBO trajectories because encounter velocities are 30× lower.

Fernández and Brunini (2000) simulate the formation of the Oort cloud within an open cluster having parameters similar to those cited above. They find that passing stars create an "inner Oort cloud" of objects having $35 \leq q(\text{AU}) \leq 1000$, $300 \leq a(\text{AU}) \leq 10^4$, $\langle e \rangle \sim 0.8$, and $\langle i^2 \rangle^{1/2} \sim 1$. Sedna may be the first discovered member of this inner Oort cloud (*Brown et al.,* 2004). Such objects coagulate in the vicinity of the giant planets and are scattered first by them. Since a scattering event changes velocities more effectively than it does positions, objects' perihelia remain at heliocentric distances of ~5–30 AU while aphelia diffuse outward. Aphelia grow so distant that objects are scattered next by cluster stars. These stars raise objects' perihelia beyond the reach of the giant planets.

We confirm the ability of cluster stars to raise the perihelion of Sedna with an order-of-magnitude calculation. During the open cluster phase, the number of stars that pass within distance q_* of the Sun is

$$N_* \sim 1 \left(\frac{q_*}{4000 \text{ AU}} \right)^2 \left(\frac{n_*}{4 \text{ pc}^{-3}} \frac{\langle v_*^2 \rangle^{1/2}}{1 \text{ km s}^{-1}} \frac{t}{200 \text{ m.y.}} \right) \quad (25)$$

A star of mass M_* having perihelion distance q_* much greater than a planetesimal's aphelion distance ($Q \approx 2a$) perturbs that object's specific angular momentum by

$$\delta h = \pm C \frac{GM_*}{\langle v_*^2 \rangle^{1/2}} \left(\frac{a}{q_*} \right)^2 \quad (26)$$

where the numerical coefficient C depends on the encounter geometry (*Yabushita,* 1972). We can derive the form of equation (26) by noting that $\delta h \sim Q \delta v$, where δv is the perturbation to the object's velocity relative to the Sun. We write δv as the tidal acceleration $GM_* Q / q_*^3$ induced by the star, multiplied by the duration $q_* / \langle v_*^2 \rangle^{1/2}$ of the encounter, to arrive at equation (26). For highly eccentric orbits $\delta q = h \delta h / (GM_\odot)$, whence

$$\frac{\delta q}{q} \sim \pm C \frac{M_*}{M_\odot} \left(\frac{a}{q_*} \right)^2 \left(\frac{2GM_\odot}{q \langle v_*^2 \rangle} \right)^{1/2} \quad (27)$$

For $M_* = M_\odot$, $q_* = 4000$ AU, $C \approx 6$ [see equation (3.17) of *Yabushita* (1972)], $\langle v_*^2 \rangle^{1/2} = 1$ km s^{-1}, and preencounter values of q = 35 AU and a = 600 AU, $\delta q / q \sim \pm 1$. Thus, Sedna's perihelion could have doubled to near its current value, $q \approx 76$ AU, by a single slow-moving cluster star. Multiple encounters at larger q_* cause q to random walk and change its value less effectively: $\langle (\delta q)^2 \rangle^{1/2} \propto (N_*)^{1/2} q_*^{-2} \propto q_*^{-1}$.

Had we performed this calculation for parameters appropriate to the present-day stellar environment, we would have found $\delta q / q \approx \pm 0.2$. The reduction in efficacy is due to the larger $\langle v_*^2 \rangle$ today.

The cluster properties cited above are averaged over a half-light radius of 2 pc (*Binney and Tremaine,* 1987). For comparison, the Hyades cluster has 4× lower n_*, 3× lower $\langle v_*^2 \rangle^{1/2}$, and 6× longer lifetime t (*Binney and Merrifield,* 1998; *Perryman et al.,* 1998); the Hyades therefore generates 2× fewer encounters than does our canonical cluster. Younger clusters like the Orion Trapezium maintain 15× higher n_* and similar $\langle v_*^2 \rangle^{1/2}$ over 200× shorter t (*Hillenbrand and Hartmann,* 1998), and therefore yield even fewer

encounters. Scenarios that invoke stellar encounters for which $q_* \ll 1000$ AU to explain such features as the edge of the classical belt require that the Sun have resided in a cluster having atypical properties, i.e., dissimilar from those of the Orion Trapezium, the Hyades, and all open clusters documented by *Binney and Merrifield* (1998). That parent bodies in extrasolar debris disks also do not extend beyond ~40–100 AU (section 3.1.2) argues against explanations that rely on unusually dense environments.

3.8. Coagulation of Neptune Trojans

Planetesimal collisions that occur near Neptune's Lagrange points insert debris into 1:1 resonance. This debris can coagulate into larger bodies. The problem of accretion in the Trojan resonance is akin to the standard problem of planet formation, transplanted from a star-centered disk to a disk centered on the Lagrange point. As with other kinds of transplant operations, there are complications: Additional timescales not present in the standard problem, such as the libration period T_{lib} about the Lagrange point, require juggling. *Chiang and Lithwick* (2005, hereafter *CL05*) account for these complications to conclude that QR-sized Trojans may form as miniature oligarchs, each dominating its own tadpole-shaped annulus in the ancient Trojan subdisk. Alternative formation scenarios for Trojans such as pull-down capture and direct collisional emplacement of QR-sized objects into resonance are considered by *CL05* and deemed unlikely. Also, the mechanism proposed by *Morbidelli et al.* (2005) to capture Jupiter Trojans cannot be applied to Neptune Trojans since Uranus and Neptune today lie inside their 1:2 MMR and therefore could not have divergently migrated across it (A. Morbidelli, personal communication). We focus on *in situ* accretion, but acknowledge that a collisionless capture scenario might still be feasible and even favored by late-breaking data; see the end of this subsection.

In the theory of oligarchic planet formation (e.g., *G04*), each annulus is on the order of 5 R_H in radial width; the number of QR-sized oligarchs that can be fitted into the tadpole libration region is

$$N_{Trojan} \sim \frac{(8M_N/3M_\odot)^{1/2} a_N}{5\,R_H} \sim 20 \qquad (28)$$

attractively close to the number of QR-sized Neptune Trojans inferred to exist today (section 2.2.3). The numerator in equation (28) equals the maximum width of the 1:1 MMR, $a_N \approx 30$ AU is Neptune's current semimajor axis, $R_H = R/\alpha$ is the Trojan's Hill radius, and $R \approx 90$ km is the radius of QR.

The input parameters of the coagulation model are the surface density σ and sizes s of small bodies in 1:1 resonance. Big bodies grow by consuming small bodies, but growth is limited because small bodies diffuse out of resonance by colliding with other small bodies. The time for a small body to random walk out of the Trojan subdisk is

$$t_{esc} \sim \frac{\rho s}{\sigma\Omega}\left[\frac{(M_N/M_\odot)^{1/2} a_N}{u/\Omega}\right]^2 \qquad (29)$$

The term in square brackets follows from noting that a small body shifts its orbital guiding center by of order its epicyclic amplitude $\sim \pm u/\Omega$ every time it collides with another small body in an optically thin disk. To escape resonance, the small body must random walk the maximum libration width. We equate t_{esc} to the growth time of a big body t_{acc} (equation (4)) to solve for the maximum size to which a large body coagulates

$$R = R_{final} \sim 100\left(\frac{2}{u/v_H}\right)^{4/3}\left(\frac{s}{20\text{ cm}}\right)^{1/3} \text{km} \qquad (30)$$

Our normalization of $u/v_H \approx 2$ is derived from $s \sim 20$ cm and $\sigma \sim 4 \times 10^{-4}$ g cm^{-2} ~ 10× the surface density inferred in QR-sized objects today; we derive u/v_H by balancing gravitational stirring by big bodies with damping by inelastic collisions between small bodies (*CL05*). For these parameter values, $t_{esc} \sim t_{acc} \sim 1 \times 10^9$ yr. Unlike Neptune-sized oligarchs that may have been ejected out of the solar system (section 3.4), all ~10–30 Trojan oligarchs in a single cloud should be present and eventually accounted for.

As speculated by *CL05*, orbital inclinations of Trojans with respect to Neptune's orbit plane might be small; perhaps $\langle i^2 \rangle^{1/2} \lesssim 10°$. A thin disk of Neptune Trojans would contrast with the thick disks occupied by Jupiter Trojans, main-belt asteroids, and nonclassical KBOs, and would reflect a collisional, dissipative birth environment. Three other Neptune Trojans have since been announced after the discovery of QR, having inclinations of 1.4°, 25.1°, and 5.3° (*Sheppard and Trujillo,* 2006). If a large fraction of Neptune Trojans have high i, we might look to the ν_{18} secular resonance, unmodeled by *CL05*, to amplify inclinations. See also *Tsiganis et al.* (2005), who find that Neptune Trojans can be captured collisionlessly; the capture process is related to "freeze-in" as described in section 3.6.4.

3.9. Collisional Comminution

Over the last few billion years, sufficiently small and numerous bodies in the Kuiper belt suffer collisional attrition. As interpreted by *Pan and Sari* (2005, hereafter *PS05*), the break in the size distribution of KBOs at $R \approx 50$ km as measured by *Bernstein et al.* (2004) (section 2.2.1) divides the collisional spectrum at small R from the primordial coagulation spectrum at large R. For the remainder of this subsection, we do not distinguish between the various dynamical classes but instead analyze all KBOs together as a single group. At $R > R_{break}$, the size spectrum $dN/dR \propto R^{-\tilde{q}_0}$, where dN is the number of objects per unit face-on area of the belt having sizes between R and R + dR (the

differential surface number density). The slope $\tilde{q}_0 \sim 5$ (see section 2.2.1 for more precise values) presumably represents the unadulterated outcome of coagulation. Bodies at this large-R end of the spectrum are insufficiently numerous to collide among themselves and undergo attrition. At $R < R_{break}$, $dN/dR \propto R^{-\tilde{q}}$, where $\tilde{q}$ derives from a quasisteady collisional cascade (*Dohnanyi,* 1969; *PS05*). By definition of R_{break}, the time for a body of radius R_{break} to be catastrophically dispersed equals the time elapsed

$$\frac{1}{N_{proj} \times \pi R_{break}^2 \times \Omega} \sim t \tag{31}$$

where πR_{break}^2 is the collision cross-section and N_{proj} is the surface number density of projectiles that are just large enough to disperse R_{break}-sized targets (catastrophic dispersal implies that the mass of the largest postcollision fragment is no greater than half the mass of the original target and that collision fragments disperse without gravitational reassembly). This expression is valid for the same assumptions underlying equation (3), i.e., for today's dynamically hot belt.

We proceed to estimate R_{break} given the parameters of the present-day Kuiper belt. For $R > R_{break}$, $N = N_0(R/R_0)^{1-\tilde{q}_0}$, where N is the surface number density of objects having sizes between R and 2 R. We estimate that for fiducial radius $R_0 = 100$ km, $N_0 \approx 20$ AU^{-2} at $a \approx 43$ AU. The minimum radius R_{proj} of the projectile that can catastrophically disperse a target of radius R_{break} is given by

$$\frac{1}{2} R_{proj}^3 v_{rel}^2 = R_{break}^3 Q^* \tag{32}$$

where

$$Q^* = Q_0^* \left(\frac{R}{R_0} \right)^y \tag{33}$$

is the collisional specific energy (*Greenberg et al.,* 1978; *Fujiwara et al.,* 1989) and v_{rel} is the relative collision velocity. Since for $R < R_{break}$ as much mass is ground into every logarithmic interval in R as is ground out (e.g., *PS05*),

$$\tilde{q} = \frac{21 + y}{6 + y} \tag{34}$$

We assume (and can check afterward) that $R_{proj} < R_{break} < R_0$ to write

$$N_{proj} = N_0 \left(\frac{R_{break}}{R_0} \right)^{1-\tilde{q}_0} \left(\frac{R_{proj}}{R_{break}} \right)^{1-\tilde{q}} \tag{35}$$

Combining the above relations yields

$$\frac{R_{break}}{R_0} \sim (\pi N_0 R_0^2 \Omega t)^{z_1} \left(\frac{v_{rel}^2}{2Q_0^*} \right)^{z_2} \tag{36}$$

where $z_1 = (6 + y)/[5y + (6 + y)(\tilde{q}_0 - 3)]$ and $z_2 = 5/[5y + (6 + y)(\tilde{q}_0 - 3)]$. For targets held together by self-gravity, $Q^* \approx 3v_{esc}^2/10$ and $y = 2$. If we insert these values into equation (36), together with $v_{rel} = 1$ km s^{-1}, $\tilde{q}_0 = 5$, $\Omega = 2\pi/(300$ yr), and $t = 3 \times 10^9$ yr, we find that $R_{break} \approx 0.4\, R_0 \approx 40$ km, in good agreement with the observed break in the luminosity function (Fig. 5) (*PS05*). The small-R end of the KBO size spectrum as observed today reflects the catastrophic comminution of bodies that derive their strength from self-gravity ("rubble piles"). Furthermore, the Kuiper belt has been dynamically hot for the last few billion years (*PS05*).

4. DIRECTIONS FOR FUTURE WORK

1. *Collisional vs. collisionless:* Most explorations of planetary migration and of how the Kuiper belt was stirred utilize collisionless gravitational simulations. But the overwhelming bulk of the primordial mass may have resided in small, collisional bodies. Simultaneously accounting for collisions and gravity might revolutionize our understanding of the clean-up (a.k.a. missing-mass) problem. Insights from the study of planetary rings will be helpful.

2. *Classical Kuiper belt object colors vs. heliocentric distance:* Do classical KBOs exhibit a trend in color from neutral to red with increasing heliocentric distance d? The two neutral classicals at $d \approx 38$ AU, contrasted with the predominantly red classicals at $d \approx 42$ AU, suggest the answer is yes (Figs. 3 and 4). Confirmation would support ideas that classicals coagulated *in situ*, and that neutrally colored resonant/scattered KBOs coagulated from small d and were transported outward. We must also ask why trends in color with birth distance d would exist in the first place.

3. *Formation of the scattered belt by Neptune-mass oligarchs:* We argue that Neptune's migration and the concomitant sweeping of secular resonances do not populate the scattered and hot classical belts with enough objects to explain observations. When account is made of the primordial size distribution of planetesimals — a distribution that should be preserved today at large sizes (sections 1 and 3.9) — the expected population of scattered/hot classical objects having sizes above 100 km is less than that observed by a factor of 50–150. We propose instead that planetesimals were deflected onto scattered/hot classical orbits by simple close encounters with marauding Neptune-mass oligarchs that have since been ejected from the solar system, and by Neptune while its orbit circularized by dynamical friction. These contentions are supported by order-of-magnitude estimates but require numerical simulations to verify.

4. *Kuiper Cliff:* Why do planetesimal disks have sharp outer edges?

5. *Binaries*: Kuiper belt binaries might prove the most informative witnesses we have to the history of transneptunian space. They hearken back to a primordially dense and cold disk in which collisions and multiple-body encounters were orders of magnitude more frequent than they are today. Binary orbit properties must also reflect how the Kuiper belt was stirred as a whole. How binary inclinations, eccentricities, and component mass ratios are distributed, and how/why the incidence of binarity correlates with dynamical class, are open issues for observer and theorist alike.

Acknowledgments. This work was supported by the Alfred P. Sloan Foundation, the National Science Foundation, and NASA. We acknowledge helpful exchanges with G. Bernstein, P. Goldreich, R. Gomes, D. Jewitt, S. Kenyon, A. Morbidelli, D. Nesvorný, R. Sari, L. Strubbe, A. Youdin, and an anonymous referee. We are grateful to the Deep Ecliptic Survey (DES) team for their unstinting support.

REFERENCES

Bernstein G. M., Trilling D. E., Allen R. L., Brown M. E., Holman M., and Malhotra R. (B04) (2004) *Astron. J., 128,* 1364–1390.

Binney J. and Merrifield M. (1998) *Galactic Astronomy,* pp. 377–386. Princeton Univ., Princeton.

Binney J. and Tremaine S. (1987) *Galactic Dynamics,* pp. 26, 440–443. Princeton Univ., Princeton.

Boehnhardt H., Tozzi G. P., Birkle K., Hainaut O., Sekiguchi T., et al. (2001) *Astron. Astrophys., 378,* 653–667.

Brown M. E., Trujillo C., and Rabinowitz D. (2004) *Astrophys. J., 617,* 645–649.

Brown M. E., Trujillo C. A., and Rabinowitz D. L. (2005a) *Astrophys. J., 635,* L97–L100.

Brown M. E., van Dam M. A., Bouchez A. H., Le Mignant D., Campbell R. D., et al. (2005b) *Astrophys. J., 632,* L45–L48.

Buie M. W., Grundy W. M., Young E. F., Young L. A., and Stern S. A. (2006) *Astron. J., 132,* 290–298.

Canup R. M. (2005) *Science, 307,* 546–550.

Chiang E. I. and Jordan A. B. (2002) *Astron. J., 124,* 3430.

Chiang E. I. and Lithwick Y. (CL05) (2005) *Astrophys. J., 628,* 520–532.

Chiang E. I., Jordan A. B., Millis R. L., Buie M. W., Wasserman L. H., et al. (C03) (2003) *Astron. J., 126,* 430–443.

Christy J. W. and Harrington R. S. (1978) *Astron. J., 83,* 1005–1008.

Dermott S. F., Malhotra R., and Murray C. D. (1988) *Icarus, 76,* 295–334.

Dohnanyi J. W. (1969) *J. Geophys. Res., 74,* 2531–2554.

Duncan M. J., Levison H. F., and Budd S. M. (D95) (1995) *Astron. J., 110,* 3073–3081.

Elliot J. L., Kern S. D., Clancy K. B., Gulbis A. A. S., Millis R. L., et al. (E05) (2005) *Astron. J., 129,* 1117–1162.

Fernández J. A. and Brunini A. (2000) *Icarus, 145,* 580–590.

Fernández J. A. and Ip W.-H. (1984) *Icarus, 58,* 109–120.

Fujiwara A., Cerroni P., Davis D., Ryan E., and di Martino M. (1989) In *Asteroids II* (R. P. Binzel et al., eds.), pp. 240–265. Univ. of Arizona, Tucson.

Funato Y., Makino J., Hut P., Kokubo E., and Kinoshita D. (2004) *Nature, 427,* 518–520.

Garaud P. and Lin D. N. C. (2004) *Astrophys. J., 608,* 1050–1075.

Gladman B., Holman M., Grav T., Kavelaars J., Nicholson P., et al. (2002) *Icarus, 157,* 269–279.

Goldreich P. (1965) *Mon. Not. R. Astron. Soc., 130,* 159–181.

Goldreich P., Lithwick Y., and Sari R. (G02) (2002) *Nature, 420,* 643–646.

Goldreich P., Lithwick Y., and Sari R. (G04) (2004a) *Ann. Rev. Astron. Astrophys., 42,* 549–601.

Goldreich P., Lithwick Y., and Sari R. (2004b) *Astrophys. J., 614,* 497–507.

Gomes R. S. (2003a) *Icarus, 161,* 404–418.

Gomes R. S. (2003b) *Earth Moon Planets, 92,* 29–42.

Gomes R. S., Morbidelli A., and Levison H. F. (2004) *Icarus, 170,* 492–507.

Gomes R. S., Gallardo T., Fernández J. A., and Brunini A. (2005) *Cel. Mech. Dyn. Astron., 91,* 109–129.

Gómez G. C. and Ostriker E. C. (2005) *Astrophys. J., 630,* 1093–1106.

Greenberg R., Hartmann W. K., Chapman C. R., and Wacker J. F. (1978) *Icarus, 35,* 1–26.

Greenberg R., Bottke W. F., Carusi A., and Valsecchi G. B. (1991) *Icarus, 94,* 98–111.

Hahn J. M. and Malhotra R. (1999) *Astron. J., 117,* 3041–3053.

Hahn J. M. and Malhotra R. (2005) *Astron. J., 130,* 2392–2414.

Hillenbrand L. A. and Hartmann L. W. (1998) *Astrophys. J., 492,* 540–553.

Holman M. J. and Wisdom J. (HW93) (1993) *Astron. J., 105,* 1987–1999.

Ida S. and Makino J. (1993) *Icarus, 106,* 210–217.

Jewitt D. and Luu J. (1993) *Nature, 362,* 730–732.

Jewitt D. and Luu J. (2000) In *Protostars and Planets IV* (V. Mannings et al., eds.), pp. 1201–1229. Univ. of Arizona, Tucson.

Jewitt D., Luu J., and Trujillo C. (1998) *Astron. J., 115,* 2125–2135.

Kalas P. and Jewitt D. (1995) *Astron. J., 110,* 794–804.

Kenyon S. J. (2002) *Publ. Astron. Soc. Pac., 793,* 265–283.

Kenyon S. J. and Bromley B. C. (2006) *Astron. J., 131,* 1837–1850.

Kenyon S. J. and Luu J. X. (1998) *Astron. J., 115,* 2136–2160.

Kenyon S. J. and Luu J. X. (1999) *Astron. J., 118,* 1101–1119.

Knežević Z., Milani A., Farinella P., Froeschle Ch., and Froeschle Cl. (1991) *Icarus, 93,* 315–330.

Krist J. E., Ardila D. R., Golimowski D. A., Clampin M., Ford H. C., et al. (2005) *Astron. J., 129,* 1008–1017.

Levison H. F. and Morbidelli A. (2003) *Nature, 426,* 419–421.

Malhotra R. (1993) *Nature, 365,* 819–821.

Malhotra R. (1995) *Astron. J., 110,* 420–429.

Matsuyama I., Johnston D., and Hartmann L. (2003) *Astrophys. J., 582,* 893–904.

Millis R. L., Buie M. W., Wasserman L. H., Elliot J. L., Kern S. D., and Wagner R. M. (2002) *Astron. J., 123,* 2083–2109.

Morbidelli A., Jacob C., and Petit J.-M. (2002) *Icarus, 157,* 241–248.

Morbidelli A., Levison H. F., Tsiganis K., and Gomes R. (2005) *Nature, 435,* 462–465.

Murray C. D. and Dermott S. F. (1999) In *Solar System Dynamics,* pp. 63–128, 225–406. Cambridge Univ., Cambridge.

Murray-Clay R. A. and Chiang E. I. (MC05) (2005) *Astrophys. J., 619,* 623–638.

Murray-Clay R. A. and Chiang E. I. (MC06) (2006) *Astrophys. J.,* in press (astro-ph/0607203).

Nesvorný D. and Dones L. (2002) *Icarus, 160,* 271–288.

Noll K. (2003) *Earth Moon Planets, 92,* 395–407.

Pan M. and Sari R. (2004) *Astron. J., 128,* 1418–1429.

Pan M. and Sari R. (PS05) (2005) *Icarus, 173,* 342–348.

Peale S. J. (1986) In *Satellites* (J. A. Burns and M. S. Matthews, eds.), pp. 159–223. Univ. of Arizona, Tucson.

Peixinho N., Boehnhardt H., Belskaya I., Doressoundiram A., Barucci M. A., and Delsanti A. (2004) *Icarus, 170,* 153–166.

Perryman M. A. C., Brown A. G. A., Lebreton Y., Gomez A., Turon C., et al. (1998) *Astron. Astrophys., 331,* 81–120.

Press W. H., Teukolsky S. A., Vetterling W. T., and Flannery B. P. (1992) In *Numerical Recipes in C: The Art of Scientific Computing,* pp. 609–639. Cambridge Univ., Cambridge.

Sheppard S. S. and Jewitt D. (2004) *Astron. J., 127,* 3023–3033.

Sheppard S. S. and Trujillo C. A. (2006) *Science, 313,* 511–514.

Stephens D. C. and Noll K. S. (2006) *Astron. J., 131,* 1142–1148.

Stern S. A., Weaver H. A., Steffl A. J., Mutchler M. J., Merline W. J., et al. (2006) *Nature, 439,* 946–948.

Strubbe L. E. and Chiang E. I. (2006) *Astrophys. J., 648,* 652–665.

Thommes E. W., Duncan M. J., and Levison H. F. (1999) *Nature, 402,* 635–638.
Thommes E. W., Duncan M. J., and Levison H. F. (2002) *Astron. J., 123,* 2862–2883.
Tiscareno M. S. and Malhotra R. (2003) *Astron. J., 126,* 3122–3131.
Trujillo C. A. and Brown M. E. (2001) *Astrophys. J., 554,* L95–L98.
Trujillo C. A. and Brown M. E. (2002) *Astrophys. J., 566,* L125–L128.
Tsiganis K., Gomes R., Morbidelli A., and Levison H. F. (2005) *Nature, 435,* 459–461.
Veillet C., Parker J. W., Griffin I., Marsden B., Doressoudiram A., et al. (2002) *Nature, 416,* 711–713.
Weaver H. A., Stern S. A., Mutchler M. J., Steffl A. J., Buie M. W., et al. (2006) *Nature, 439,* 943–945.
Weidenschilling S. J. (2002) *Icarus, 160,* 212–215.
Yabushita S. (1972) *Astron. Astrophys., 16,* 395–403.
Youdin A. N. and Shu F. H. (2002) *Astrophys. J., 580,* 494–505.
Youdin A. N. and Chiang E. I. (2004) *Astrophys. J., 601,* 1109–1119.
Youdin A. N. and Goodman J. (2005) *Astrophys. J., 620,* 459–469.