Using the Main Asteroid Belt to Constrain Planetesimal and Planet Formation

William F. Bottke
Southwest Research Institute and the Institute for the Science of Exploration Targets (ISET), Boulder, CO

Alessandro Morbidelli
Dep. Lagrange, CNRS, Observatoire de la Cote d'Azur, Universite de Nice Sophia-Antipolis; Nice, France

3.1 Introduction

The largest population of remnant planetesimals still found in the inner solar system is the main asteroid belt, located between 2.1-3.2 AU. According to meteorite studies, many of the largest bodies there formed $\sim 4.56Ga$ during planet formation processes. Since that time, however, they have been subject to collisional and dynamical evolution. A key goal of asteroid belt studies, therefore, is to turn back the clock, understand how the main belt has changed, and use the information to infer to original properties of primordial main belt planetesimals. By doing so, we not only provide powerful constraints on planetesimal and planet formation models, but we can also place meteorite constraints into their appropriate solar system context.

A problem in interpreting what we know about asteroids, however, is that collisonal and dynamical evolution are coupled to one other. For example, assuming a given dynamical excitation state for a small body population, more collisional evolution takes place when a population is large than when it is small. Thus, if dynamical effects suddenly remove bodies from a population, disruption and cratering events must drop as well. Similar, a population with low eccentricities and inclinations will undergo little collisional grinding, while one with large values will grind much faster.

For this reason, our discussion starts with what has been inferred about the collisional evolution of the asteroid belt. This leads into how it has been affected dynamically by the processes that led to the origin of our planets.

3.2 Constraints for collisional evolution

Given the enormous number of possibilities that can take place in collisional evolution models for given assumptions, and the importance of dynamical excitation, dynamical removal, and stochastic breakup events, it is critical that planetesimal formation and collisional/dynamical evolution models be tested against as many constraints as possible. This potentially allows us to rule out certain scenarios and place higher degrees of confidence in successful solutions. Here we discuss many of the constraints that need to be considered when modeling the collisional evolution of the main belt.

3.2.1 Wavy main belt size frequency distribution

A key constraint comes from the main belt size frequency distribution (SFD). Reasonable estimates of this SFD can be found in several places, and reviews of this topic can be found in Jedicke et al. (2002), Mainzer et al. (2015), and Masiero et al. (2015). Here we turn the absolute magnitude H distribution described in Jedicke et al (2002) into a size distribution using the relationship between asteroid diameter D, absolute magnitude H, and visual geometric albedo p_v provided by Fowler and Chillemi (1992):

$$D = \frac{1329}{\sqrt{p_v}} 10^{-H/5}. \quad (1)$$

By setting the $p_v = 0.092$, we get the population is shown in Fig. 3.1 (see Bottke et al. 2005 for details). The observed and debiased main belt SFD is wavy, with "bumps" near diameter $D \sim 3$ km and one near $D \sim 100$ km. The latter bump is at a similar size to the one seen within the transneptunian objects SFD (e.g., Fraser et al. 2014), and it is suggestive of something fundamental about this size and planetesimal formation. Additional discussion of the origin and nature of the shape of this size distribution can be found in Cuzzi et al. (2010) and Johansen et al. (2015). Note that even more precise main belt constraints can be obtained by treating different regions of the main belt separately; see Cibulková et al. (2014) for details.

3.2.2 Asteroid families

Asteroid families are remnants of cratering and catastophic disruption events in the main belt. Identified by their clustered values of proper semimajor axes $a_{\rm p}$, eccentricities $e_{\rm p}$, and inclinations $i_{\rm p}$ (see Nesvorny et al. 2015), they can be used to deduce reasonable scaling relationship that describe how projectiles catastrophically disrupt large asteroids. One must use caution in applying asteroid families as model constraints, however, because (i) smaller families can potentially be eliminated over time by collisional and dynamical processes and (ii) estimates of ancient family ages, made using dynamical methods (e.g., Vokrouhlicky et al. 2006), have uncertainties. This has led groups like Bottke et al. (2005a, b) to use very large families as constraints, specifically families whose parent body was large enough that family fragments could not be easily erased over 3.5-4 Gyr of evolution. We have very little knowledge of families older than 3.5-4 Ga to date, and it plausible their clusters were scattered by large scale dynamical processes (see below).

Using numerical hydrocode simulations that track how asteroids likely undergo disruptions, Durda et al. (2007)

argued that approximately 20 observed families created by catastrophic disruptions of parent bodies with sizes $D_{\rm PB} > 100\,{\rm km}$, where the ratio of the largest fragment's mass to the parent body mass is $M_{\rm LR}/M_{\rm PB} < 0.5$ (Fig. **??**). Specifically, they argued that the parent body disruptions over the last 3.5 Ga occurred within incremental logarithmic-separated bins centered on diameters $D = 123.5, 155.5, 195.7, 246.4, 310.2$, and 390.5 km were 5, 5, 5, 1, 1, 1, respectively. Note that these values do not include large cratering events, such as the Vesta family.

A recent update to these family estimates can be found in Cibulková et al. (2014). Their values are generally similar, but there are subtleties; see Bottke et al. (2015) for a discussion. For example, it is possible that a few small families in the main belt today are remnants, or "ghosts", of much larger older families (e.g., possibly (918) Itha; Brož et al. 2013). We define a ghost family here as one so ancient that collisional and dynamical effects have rendered it nearly unrecognizable to standard identification methods (e.g., loss of numerous smaller members, sufficient orbital element spreading that a family cluster is hard to identify, etc.). We believe that at least a few ghost families exist, but a smoking gun for them has yet to be identified. If it can be demonstrated that numerous ghost families still reside in the main belt population, the nature of our proposed asteroid family constraint will change substantially. For example, numerous missing families could suggest that our disruption scaling relationships used for asteroid breakup events need to undergo revisions.

3.2.3 Impact basins on (4) Vesta

Vesta is the second largest asteroid in the main belt, with a diameter of 525 km. It is differentiated and has a large intact basaltic crust. Recently visited by the Dawn spacecraft, Vesta can be considered a primordial "witness plate" for the bombardment history of the asteroid belt. Moreover, considerable knowledge of Vesta comes from eucrites, howardites, and diogenite meteorites that are thought to come from Vesta. They tell us Vesta's crust was put in place shortly after Vesta differentiated, approximately 2-3 Myr after the formation of the calcium-aluminum inclusions CAIs (Russell et al. 2015). The intact nature also places hard limits on how much collisional grinding could have ever taken place in the main belt (e.g., Davis et al. 2002).

Even better constraints than the Vesta crust, however, are two enormous impact basins that dominate its southern hemisphere: Rheasilvia, a 505 km diameter crater with an estimated crater retention age of 1 Gyr, and Venenia, a 395 km crater with a crater retention age of > 2 Gyr (Marchi et al. 2012). Rheasilvia, being younger, overlaps with and has largely obscured Veneneia (Schenk et al. 2012; Jaumann et al. 2012). These basins likely produced the majority of the observed Vesta family, a spread out swarm of $D < 10$ km asteroids in the inner main belt with inclinations and spectral properties similar to Vesta itself. They also produced a set of fracture-like troughs, or graben, for each basin. Rheasilvia's are located near the equator and lie along a plane that is orthogonal to the basin center. Veneneia's are similar in character but are oriented to be orthogonal to its basin center (Buczkowski et al. 2012).

Vesta shows no obvious signs that basins similar in size to Rheasilvia or Veneneia were ever erased or buried after its basaltic crust was put in place; nothing notable is detected in Vesta's topography, and there are no unaccounted sets of troughs that could be linked with a missing or erased basin. This means Vesta is probably complete in Rheasilvia- or Veneneia-sized basins. Thus, the size of many primordial populations as well as how long they could have lasted on Vesta-crossing orbits (e.g., main belt asteroids, leftovers planetesimals from terrestrial and giant planet formation, the putative late heavy bombardment population, Jupiter-family comets, etc.), are constrained by the fact that Vesta does not have > 2 such basins.

3.2.4 Near-Earth asteroids and lunar craters

Most of our Solar System's near-Earth asteroids (NEAs) are thought to come from the asteroid belt, with the bodies drifting into resonant "escape hatches" by the non-gravitational (thermal) Yarkovsky/YORP effects (see Vokrouhlický et al. 2015). Some of these bodies have gone on to hit ancient worlds like the Moon. Accordingly, both the observed planet-crossing asteroid populations and the crater populations found on the Moon (and other bodies) can be used to constrain the approximate nature of the past and present main belt SFD.

A reasonable estimate of the NEA population is shown in Fig. 3.1. Its wavy shape is broadly similar to the main belt, though some differences exist; recall that the main belt SFD is modified en route to the NEA population by Yarkovsky-driven asteroid migration (e.g., Morbidelli and Vokrouhlický 2003). This non-gravitational thermal force is caused by sunlight and it affects the orbital motion of $D < 40$ km bodies. When these bodies heat up in the Sun, they eventually re-radiate the energy away as heat, which in turn creates a tiny thrust. This recoil acceleration is much weaker than solar and planetary gravitational forces, but it can produce substantial secular semimajor axis changes over timescales ranging from many millions to billions of years. This can allow some small main belt asteroids to drift far enough to reach a dynamical resonance capable of pushing it out of the main belt and onto a terrestrial planet-crossing orbit.

Interestingly, the shape of the NEA population in Fig. 3.1 is similar to the best available crater SFD of lunar craters formed over the last 3.2 Ga. The ages of lunar craters can be subdivided into two components; the youngest are Copernican-era, while the older ones are in the Eratosthenian-era (Fig. 3.2) (McEwen et al. 1997; Ivanov et al. 2002). Copernican craters are often considered to be < 1 Gyr old, while Eratosthenian-era craters are between 1-3.2 Ga, with the oldest age defined by the 3.2 Gyr old ages of samples returned by the Apollo 12 astronauts (Stöffler and Ryder 2001).

In Fig. 3.2, we see that not only do the combination of Copernican and Eratosthenian-era craters have the same basic shape as the NEA population, but they are also roughly a

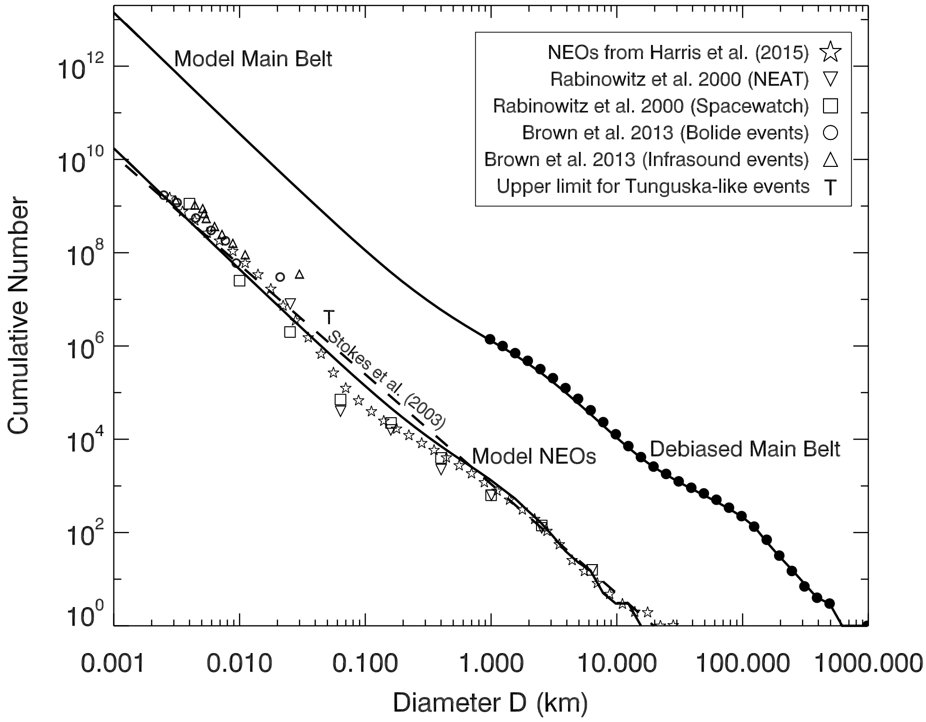

Fig. 3.1.— The estimated values of the present-day main belt and NEO populations according to Bottke et al. (2005b) model runs (solid lines). For reference, we plot our results against an estimate of the NEA population made by Stokes et al. (2003), who assumed the $D < 1$ km size distribution was a power-law extension of the $D > 1$ km size distribution, and a population discussed in Harris and D'Abramo (2015) and Harris et al. (2015). Our model main belt population provides a good match to the observed main belt (solid black dots). Most diameter $D \lesssim 100$ km bodies are fragments (or fragments of fragments) derived from a limited number of $D \gtrsim 100$ km breakups (Bottke et al. 2005a). Our NEA model population is compared to estimates derived from telescopic surveys (Rabinowitz et al. 2000) as well as satellite and infrasound detections of bolide detonations in Earth's atmosphere (Brown et al. 2013). For reference, we also include an upper limit estimate of 50 m NEAs based on the singular airblast explosion that occurred over Tunguska, Siberia in 1908. A mismatch between the NEA model and data is seen near $D \sim 0.1$ km.

factor of 3 higher than the Copernican-era craters alone. If the ages suggested above are reasonable, the simplest model would suggest the delivery of NEAs to the inner solar system and Moon over the last 3.2 Gyr has been relatively stable (to a factor of 2 or so). There are additional possibilities, such as a sizable fraction of early Eratosthenian craters coming from a different source (e.g, Bottke et al. 2012). If true, the main belt contribution to the NEA population over this interval would be lower. Regardless, we infer that main belt and NEA SFDs probably had to achieve a quasi-steady state lasting several billions of years. This allows us to rule out scenarios where a very large main belt SFDs could potentially be ground down over billions of years of comminution, with the observed SFD only achieved near the present time (see Davis et al. 2002). Such models should produce strongly-decaying lunar impact fluxes over the last 3 Gyr, and they are not observed.

3.2.5 Additional constraints

Many additional constraints can be employed to test collisional models of the asteroid belt, though we do not discuss them here for space reasons. They include (i) the population of certain types of asteroid binaries classified as SMAshed Target Satellites (SMATS) (Durda et al. 2004; 2007), (ii) the rotation rates and spin states of certain asteroids (see Bottke et al. 2015), (iii) the cosmic ray exposure ages of stony meteorites (e.g., Eugster 2003), (iv) the orbital distribution of fireballs (e.g., Morbidelli and Gladman 1998), (v) the population of V-type asteroids across the main belt (see Scott et al. 2015), (vi) the crater records found on Mercury, Venus, Earth, and Mars (e.g., Ivanov et al. 2002), (vii) smaller asteroid families not discussed here (Nesvorný et al. 2015), and (viii) the shock degassing ages of meteorites (e.g., Marchi et al. 2013).

3.3 Reconstructing the original asteroid belt

3.3.1 A brief description of generic collisional models

With these constraints in hand, we can consider modeling the collisional evolution of the main asteroid belt. In essence, models like these involve the solution of a fairly straightforward differential equation. The input is an initial SFD for the asteroid belt denoted as $N(D, t)$, with the bodies binned in logarithmic intervals as a function of diameter. The goal of the solution is to compute the time rate

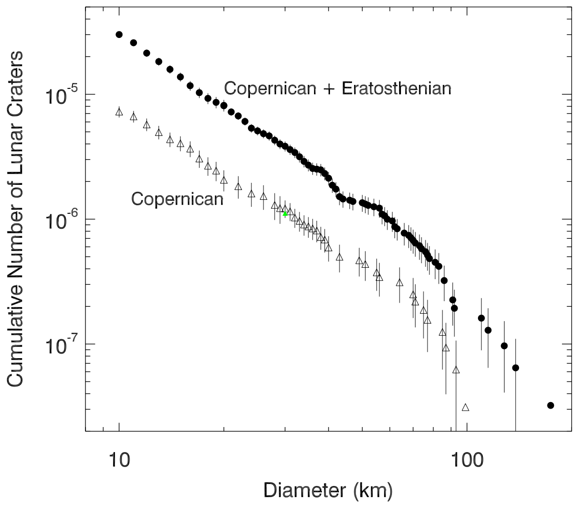

Fig. 3.2.— Lunar craters in the Copernican- and Copernican- and Eratosthenian-eras as defined by Wilhelms et al. (1978) and McEwen et al. (1997). The absolute ages of these craters are often considered < 1 and < 3.2 Gyr old (Stöffler and Ryder 2001). The plotted Copernican-era craters are a combination of nearside craters (Wilhelms 1987) and farside rayed craters (McEwen et al. 1997).

of change in the population per unit volume of space over a size range between diameter D and $D+dD$. In a schematic form, it can be written as:

$$\frac{\partial N}{\partial t}(D,t) = -I_{\rm COLL} + I_{\rm FRAG} - I_{\rm DYN}. \qquad (2)$$

Here $I_{\rm COLL}$ is the net number of bodies that leave between D and $D + dD$ per unit time from collisions (i.e., it is a "sink" for bodies in the SFD). The net number of collisions taking place at every timestep is calculated by determining how many projectiles from other size bins are capable of producing either a cratering or a catastrophic disruption event among bodies between D and $D + dD$. Note that other mass loss processes can easily be included here as well, such as the loss of material via non-gravitational YORP torques that can spin up asteroids fast enough that they shed mass.

The results of the $I_{\rm COLL}$ calculation are sent to the function $I_{\rm FRAG}$, which describes the number of bodies entering a given size bin per unit time that were produced by the fragmentation of larger bodies (i.e., it is a "source" for bodies in the SFD). This allows large asteroids act as a reservoir for smaller bodies, with collisional evolution or some other process liberating fragments over time. Finally, the equation accounts for $I_{\rm DYN}$, which is the number of bodies lost from a given size bin via dynamical processes, such as an object escaping through a dynamical resonance (i.e., it is a "sink" for bodies in the SFD). Note that $I_{\rm DYN}$ is often enacted over the entire main belt SFD, which is reasonable for global dynamical removal mechanisms like sweeping resonances or the effects of migrating planets.

The details of these functions are important as well. Some key parameters that need to be input into the codes include: (i) the collision probabilities between asteroids and different populations as well as how they might change with time, (ii) asteroid disruption scaling laws, often referred to as $Q^*_{\rm D}$ functions, which may be different for different asteroid compositions (i.e., $Q^*_{\rm D}$ is the energy per unit mass needed to send 50% of the debris away at escape velocity; Fig. 4), (iii) asteroid fragmentation laws that describe how a disrupted asteroid's mass is distributed into a new fragment SFD, (iv) a description of how main belt populations are depleted via different dynamical removal mechanisms, and (v) whether asteroid fragmentation via so-called thermal spin processes (YORP) are a major player in the destruction of small asteroids. For even more ambitious modelers, one can consider the inclusion of how planetesimals form and grow within the solar nebula, how some asteroids may have been affected by so-called "hit and run" collisions (Asphaug et al. 2015; Scott et al. 2015), and the possible inclusion of bodies of various sizes into the main belt zone by dynamical processes (Bottke et al. 2006; Levison et al. 2009; Walsh et al. 2011).

Ultimately, to model collisional evolution in the primordial asteroid belt, we need to make assumptions about the excitation of asteroid belt bodies at early times. For example, the process that caused the main belt population to become dynamically excited should have also driven many primordial main belt asteroids onto planet-crossing orbits (see below). While their orbits were short-lived, their higher eccentricity and inclinations would have allowed them to slam into the surviving main belt asteroids at high velocities for tens of Myr (e.g., Bottke et al. 2005b; Davison et al. 2013; Marchi et al. 2013). Moreover, if the primordial main belt once had considerably more mass, as discussed below, these departed bodies could be responsible for a considerable amount of collisional evolution in the main belt.

A related issue is that the primordial main belt has likely been struck by sizable but transient populations on planet-crossing orbits, such as leftover planetesimals from the terrestrial planet region (Bottke et al. 2006; 2007), ejecta from giant impacts in the terrestrial planet region (Bottke et al. 2015), comet-like planetesimals dispersed from the primordial disk during giant planet migration (Brož et al. 2013), and Jupiter-Saturn zone planetesimals pushed into the inner solar system via giant planet migration and/or evolution (Walsh et al. 2011; Turrini et al. 2011, 2012). Most of these dramatic events are thought to take place during the first 500 Myr of Solar System history. The nature and evolution of these populations is uncertain, such that dynamical models are needed to set limits on what they were plausibly like. Under certain conditions, they could also account for abundant collisional grinding in the main belt.

Given these limitations, all one can do is the best they can with what they have. This means choosing parameters that are reasonable within the bounds of what is known and then testing model results against the available constraints. The interpretation of even good matches, though, must always be met with caution. A discussion of recent advances along these lines can be found in Bottke et al. (2015b).

3.3.2 Estimating collisional evolution in the primordial main belt

A key goal in discerning how the main belt evolved concerns the initial SFD created by planetesimal formation mechanisms. Given the uncertainties surrounding the origin of the planets, a enormous range of starting SFDs are theoretically plausible. Many possibilities, however, can be ruled out by testing them against the above constraints.

For example, Bottke et al. (2005a,b) evaluated a wide range of initial SFDs and Q_D^* functions to determine which combinations work the best at reproducing the observational constraints discussed above. They found that Q_D^* functions similar to those derived in numerical SPH experiments of asteroid breakup events (Benz and Asphaug 1999) tended to work the best (Fig. 3.3), though this made their $D > 100$ km asteroids very difficult to disrupt. This led them to infer that the shape of the main belt SFD for $D > 100$ km asteroids was probably close to its primordial shape (Fig. 3.1). This prediction was similar to those made by several pioneering papers from the 1950's and 1960's (Kuiper et al. 1958; Anders 1965; Hartmann and Hartmann 1968).

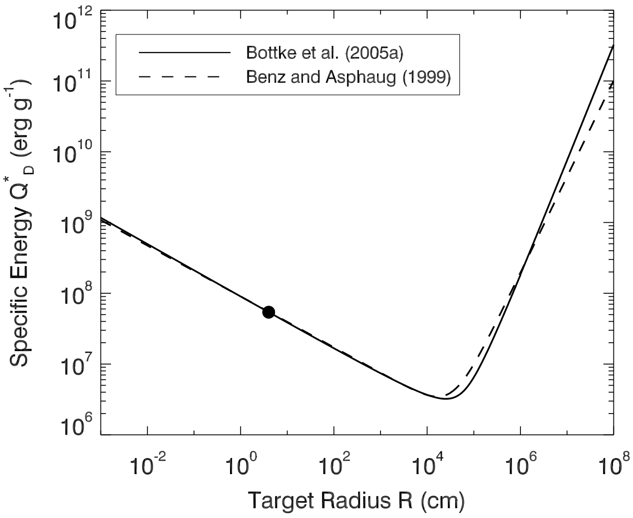

Fig. 3.3.— The critical impact specific energy $Q_{\rm D}^*$ defined by Benz and Asphaug (1999). This function is the energy per unit target mass delivered by the projectile that is required for catastrophic disruption of the target, such that one-half the mass of the target body escapes. The dashed line is the function derived by Bottke et al. (2005a) for their modeling results. Both functions pass through the normalization point $(Q_{\rm D}^*, D)$ set to $(1.5 \times 10^7$ erg g^{-1}, 8 cm), which was determined using laboratory impact experiments.

Next, Bottke et al. tested initial main belt SFDs where the incremental power law slope of -4.5 between $100 < D < 200$ km had been extended to $D < 100$ km bodies. This eliminated the observed bump near $D \sim 100$ km. They found bodies in this size range were so difficult to disrupt that initial SFDs with these shapes could not reproduce constraints. They argued from this that the bump near 100 km in the main belt SFD is primordial and that $D < 100$ km bodies probably had a shallow power law slope. Accordingly, this would indicate the planetesimal formation process favors the creation of bodies near 100 km (or larger), with smaller bodies increasingly fragments produced by the disruption of large asteroids. Examples of the starting conditions tested by Bottke et al. (20050 are found in Fig. 3.4. They found a best fit in their runs for an elbow near $D \sim$ 110-120 km. These results may act as a guide for those studying planetesimal formation processes (e.g., Morbidelli et al. 2009; Johansen et al. 2015). They also explain why a similar shape is seem among the transneptunian population.

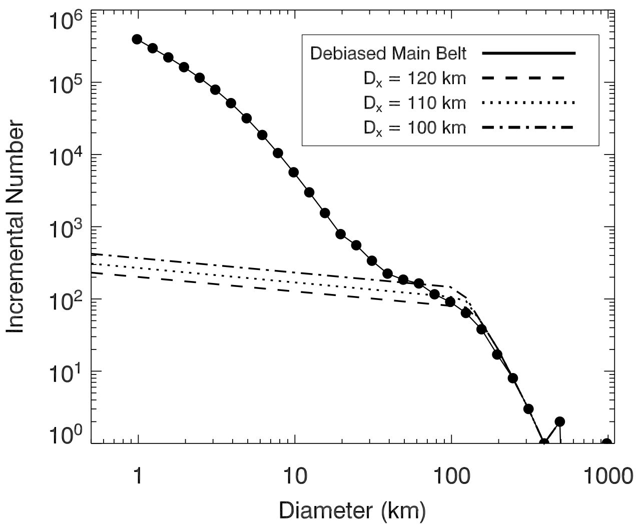

Fig. 3.4.— The debiased main belt size frequency distribution as defined in the main text (solid line). The dashed curves show possible initial shapes of the primordial main belt SFD (Bottke et al. 2005a). They found a best fit in their runs for an elbow near $D \sim$ 110-120 km. It is likely the primordial population was larger that the SFDs shown here, with most of the mass eliminated by dynamical processes.

The main belt in the successful scenarios created fragments that developed a wave-like shape (Fig. 3.5). Here a bump develops near $D \sim$ 2-3 km, which is a byproduct of the "V"-shaped $Q_{\rm D}^*$ function from Fig. 3.4. Effectively, self-gravity among $D > 200$ m objects makes them increasingly difficult to disrupt. This produces an "overabundance" of $D \sim 200$ m objects that induces a wave-like perturbation into the main belt size distribution. Impacts eventually produce a shape for its size distribution that approaches that of the observed main belt.

The next step is to try to quantify how much collisional evolution has taken place their over its history. This means choosing a starting SFD and then evaluating what it takes to reach its present-day state. The problem is there are many pathways to get from this main belt starting point to the present-day SFD, and the available constraints may be in-sufficient to tell us which ones are favored.

One way to glean insights into this issue is to adopt a simplistic but useful metric that can help us determine what different evolutionary paths might do. First, we assume that the main belt is roughly self-contained in terms of collisions, such that we can largely ignore impacts from external sources like escaped main belt asteroids, leftover planetesimals, comets, etc. Second, we assume the collision probabilities and impact velocities of asteroids hitting one another have remained unchanged over the main belt's history. Third, we assume the shape of the main belt's SFD has been close to its current shape for most of its history (Fig. 3.4), though it may have been larger in the past. We define this size to be a factor $f_{\rm MB}$, the ratio of the main belt's SFD during some past interval of time defined as ΔT over the present-day main belt SFD. Together, these values allow us to estimate the degree of collisional evolution experienced by the main belt in terms of the time exposed to different population sizes.

This metric allows to play with evolution scenarios, provided the pseudo-time is independent of the details of the dynamical depletion mechanisms. The simplest example is the nominal case where the current main belt SFD ($f_{\rm MB} = 1$) undergoes collisional evolution over its lifetime ($\Delta T = 4.56$ Gyr). The two values multiplied together yield 4.56 Gyr of collisional grinding. In a more complicated example, we assume a dynamically excited primordial main belt had $f_{\rm MB} = 300$ for 3 Myr (0.003 Gyr). At that point, most of the population was lost via escaping embryos or a migrating Jupiter, which reduced it to $f_{\rm MB} \sim 5$ for ~ 0.5 Gyr. Then, at ~ 4 Gyr, 80% of the bodies were lost via sweeping resonances driven by late giant planet migration, which left the surviving population close to its current state ($f_{\rm MB} = 1$) for the next ~ 4 Gyr. Taking all of the multiples, one can say that collectively the survivors experienced $(0.9 + 2.5 + 4) = 7.4$ Gyr of collisional evolution. This pseudo-time tells us that this main belt roughly experienced the collisional evolution equivalent of a $f_{\rm MB} = 1$ main belt going through 7.4 Gyr of comminution.

Using a collisional model that took advantage of these concepts, as well as the constraints above (e.g., shape of the main belt size distribution; number and nature of asteroid families, etc.), Bottke et al. (2005a) found median pseudo-times of 7.5-9.5 Gyr for their best fit runs, with error bars of a few Myr on each end of this range. An example of one of their runs is shown in Fig. 3.5. Their interpretation was that the main belt SFD obtained its wavy shape by going through an early time interval where the main belt survivors were exposed to many more projectiles than are observed today. This could suggest that much of the primordial main belt population was due lost to dynamical processes and/or that external impactors were effective in beating up the primordial main belt population. Either way, the wavy main belt SFD could be considered a "fossil" produced in part by early collisional evolution in the primordial main belt.

Another key property of Fig. 3.5 is that once it achieves the shape of the current main belt's SFD, it tends to keep that shape for an extended time. This would explain why the main belt SFD potentially remained in a near steady state condition for billions of years. While it would constantly changing and losing bodies by collisional, dynamical, and YORP spin up processes, it would also be steadily replenished by new large breakup events. This means the vast majority of disruption events produce too few fragments to push the main belt SFD out of equilibrium for very long. This result also explains why the crater populations on the Moon appear to have been hit by a projectile population with a similar shaped SFD for an extended period.

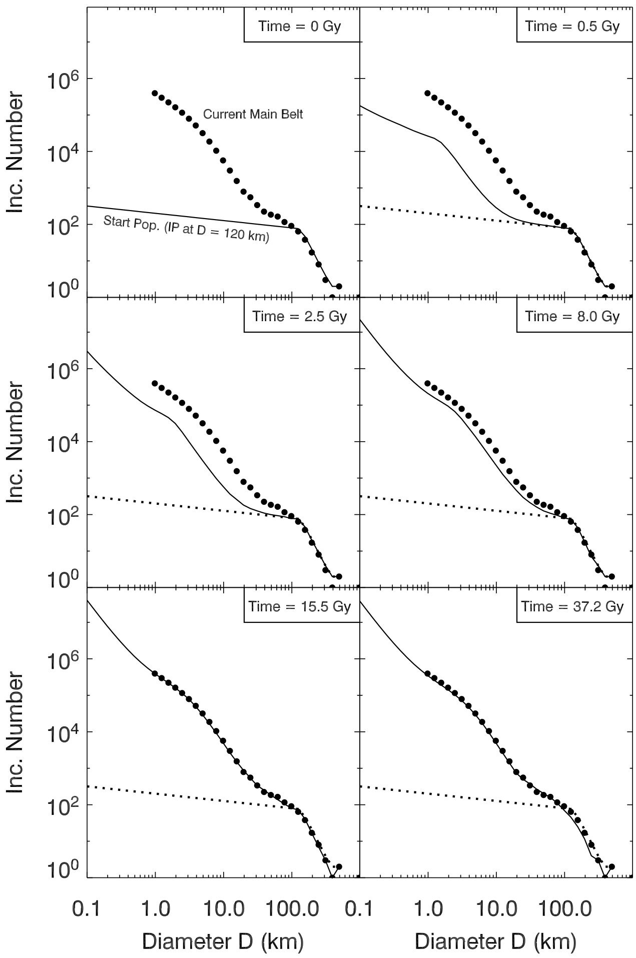

Fig. 3.5.— Six snapshots from a representative run where Bottke et al. (2005a) tracked the collisional evolution of the main belt size distribution for a pseudo-time of 50 Gy. This run uses a starting population with $D_x = 120$ km. The bump near $D \sim 120$ km is a leftover from accretion, while the bump at smaller sizes is driven by the transition at $D \sim 0.2$ km between strength and gravity-scaling regimes in $Q^*_{\rm D}$. The model main belt achieves the same approximate shape as the observed population at $t_{\rm pseudo} = 9.25$ Gyr (not shown, but it looks identical to 15.5 Gy time frame). The model closely adheres to the observed population for many Gyr after this time. Eventually, comminution eliminates enough $D > 200$ km bodies that the model diverges from the observed population.

3.4 Formation and dynamical constraints for the main belt asteroids

3.4.1 Could the asteroid belt have formed with low mass?

The classical view is that the asteroid belt had to contain about an Earth mass of material and subsequently it lost most of its mass by dynamical and/or collisional evolution. This assumption in an originally massive asteroid belt was based on two considerations.

The first consideration is that to form sizable asteroids (tens to one thousand kilometer in diameter) within the characteristic timescale of chondrite accretion (a few My; Villeneuve et al. et al., 2009; Connelly et al., 2012), the solid component of the disk required a lot of mass in smaller objects (e.g. the order of km-size planetesimals). Wetherill (1989) estimated that the initial mass had to be at least 100 $\times$ the current one. If the belt originally was low mass, collisions among small planetesimals would have been too rare and thus sizable asteroids would have grown too slowly. In principle, it is possible that the collisional coagulation process formed only a small number of large asteroids, and that most of the mass remained in small planetesimals, later removed by collisional grinding. But simulations of the collisional coagulation process (e.g. Wetherill and Stewart, 1993; Weidenschilling et al., 1997) show that most of the mass is incorporated in big objects. This suggests that a massive belt filled with sizable asteroids was a plausible outcome.

The second consideration is based on the concept of the so-called Minimum Mass Solar Nebula (MMSN: Weidenschilling, 1977; Hayashi, 1981). The MMSN is the result of an attempt to reconstruct the original mass distribution in the Solar System. It is obtained by taking the current mass incorporated in the planets, adding the mass of the missing elements required to restore a solar composition, and spreading the resulting mass into annuli that stretch between the half-way distances between neighboring planets. The resulting MMSN surface densities in the neighborhoods of Venus, the Earth and the giant planets scale approximately as $1/r^{3/2}$, where r is the heliocentric distance. However, the MMSN surface density computed from the current mass of the asteroid belt is lower by more than two orders of magnitude than that obtained by interpolation between the Venus-Earth region and the giant planet region; in the vicinity of Mars the MMSN density is more than an order of magnitude lower than that obtained by said interpolation.

One interpretation of this oddity, which we defined here as scenario 1, is that the original mass distribution in the proto-planetary disk was in fact smooth, but some mechanism depleted the Martian and asteroid belt zones of most of their solid mass. Another alternative, which we call scenario 2, is that planetesimal and/or planetary embryo formation was increasingly inefficient between the Earth and the Jupiter zone, such that most of the solid mass there did not make it into sizable bodies.

Our view of planetesimal formation, however, has radically changed over the last decade or so. In fact, the problem of forming the first putative km-sized planetesimals from dust particles has never been solved. Binary collision between dust aggregates leads to bounces and/or break-ups when particles reach sizes in the mm-cm range (mm-size barrier: Güttler et al., 2009). Moreover, meter-size boulders, even if they had formed, would have spiraled very rapidly towards the Sun by gas drag and therefore they would have been lost before having a chance to coagulate with other objects and form larger planetesimals. This is often referred to as the meter-size barrier for planetesimal formation (Weidenschilling, 1977b).

Given these problems, it was proposed that large planetesimals, perhaps around 100 km in diameter or larger, formed directly from self-gravitating clumps of small particles (see Johansen et al., 2015, for a review, and Weidenschilling et al. 2011 for an alternative view). These clumps would have formed by the interaction of the particles with the turbulent structures of the disk of gas (Johansen et al., 2007; Cuzzi et al., 2010), the particles themselves being able to generate turbulence in the disk via the Kelvin-Helmoltz instability (Johansen et al., 2006) or the streaming instability (Youdin and Goodman, 2005; Youdin and Johansen, 2007).

Given these new models, we can once again consider the mass distributions in the Mars- and asteroid belt zones. For our second scenario from above, it is conceivable that, under some conditions, only a small number of sizable bodies form in a given region, and thus they would cumulatively carry a small net mass. This could happen, for instance, if the regions are crossed by a massive flow of small particles, but the latter only sporadically manage to form self-gravitating clumps. In this case, most of the mass would just pass through the region, but no large mass would reside in the region at any given time (e.g., Levison et al. 2015a,b).

There is in fact intriguing evidence that the so-called dynamically "cold" population of the Kuiper belt formed this way. Here cold means the eccentricities and inclinations of the source population was low. It has been shown that almost all large cold KBOs are binaries (Noll et al., 2014). The best model for reproducing observations to date is through the contraction of a cloud of small particles into two orbiting self-gravitating sizable clumps (Nesvorny et al., 2010). The large fraction of binaries found there excludes the possibility that the cold population was ever scattered by proto-planets (Parker et al., 2011) or was exposed to an intense phase of collisional evolution (Petit and Mousis, 2004; Nesvorny et al., 2011). The "cold" eccentricities and inclinations of this part of the Kuiper belt also imply that both collisional and dynamical depletion are probably not an option to explain its low mass. Yet, the cold Kuiper belt contains a tiny total mass (Fraser et al., 2014). Taken together, these considerations suggest that the cold KBOs formed from the contraction of self-gravitating clumps of small particles, but the total mass of the cold KBO population was always small. Thus, at least in one part of the solar system, an observed population is in line with scenario 2.

It is unclear whether the same is true for the asteroid belt. Unlike the cold Kuiper belt, which resides at the extreme outskirts of the Solar System, the asteroid belt is bracketed between two regions where formation of massive objects (large planetesimals and protoplanets) was efficient: the terrestrial planet region and the giant planet region. This raises the question of whether accretion in the Martian and asteroid belt regions were likely to only produce only a small number of large asteroids. To do so, a drastic change in the properties of the particles in these regions had to take place. For example, perhaps the rock-ice particles drifting inward from the outer Solar System lost their ice at the snowline via sublimation just outside of the asteroid belt zone (Kretke and Lin, 2007; Levison et al., 2015a,b). This might cause the remnant to disintegrate into smaller particles that would be more difficult to accrete (Morbidelli et al., 2015). Unfortunately, our knowledge on particle coagulation and protoplanetary disk structure are still in their infancy, so we cannot yet say for certain whether scenario 2 is plausible for the Mars and asteroid belt zones. For this reason, it is useful to turn to additional constraints to explore what possibly happened.

3.4.2 Orbital excitation, radial mixing

A key characteristic of the asteroid belt population today is its orbital excitation, i.e. the fact that the eccentricities and inclinations of *many* asteroidal orbits are large (e.g, Petit et al. 2002). The median proper inclination of $D > 100$ km asteroids, most which are considered primordial planetesimals (see below), is 11 deg, while the median proper eccentricity is 0.145. Perhaps more importantly, the values of eccentricities and inclinations of these asteroids are considerably dispersed; the former ranges between 0 and 0.30 (with the limit being those that reach Mars-crossing orbits), while the latter ranges between 0 and 33 degrees. The reader should be aware that, whatever the preferred formation mechanism, planetesimals are expected to have formed on circular and co-planar orbits. Thus, one or more dynamical excitation mechanism(s) within the primordial asteroid belt were needed to stir up eccentricities and inclinations to their dispersed values. Moreover, asteroid eccentricities and inclinations do not show a strong dependence on semimajor axis, so the mechanism cannot excite one part of the asteroid belt while another part is left in a much colder state.

A second key characteristic of the asteroid belt is the partial mixing of taxonomic classes. Asteroids can be grouped into many taxonomic classes on the basis of their visual and infrared spectroscopic signatures (Tholen, 1984; Bus and Binzel, 2002; DeMeo et al., 2009). As shown first by Gradie and Tedesco (1982) for the largest asteroids, the inner belt is dominated by S-complex asteroids, many of which are probably related to the meteorites known as ordinary chondrites (Binzel et al. 1996). The central belt (2.5-3.2 AU) is dominated by C-complex asteroids, probably related to carbonaceous chondrites (Burbine et al., 2002). The Cybeles asteroids (3.2-3.7 AU), the Hilda asteroids (in the 3/2 mean motion resonance with Jupiter) and the Jupiter Trojan asteroids (in the 1/1 resonance with Jupiter) are dominated by P-and D-type asteroids. The C2 ungrouped meteorite "Tagish Lake" has been proposed to be a fragment of a D-type asteroid (Hiroi et al., 2001).

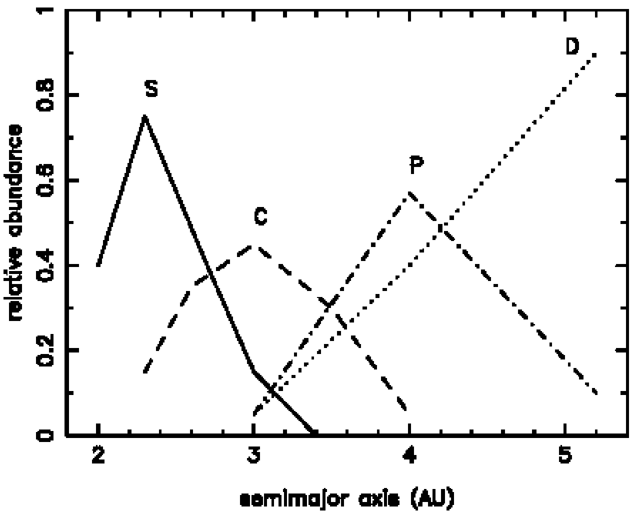

Fig. 3.6.— The relative distribution of large asteroids ($D > 50$ km) of different taxonomic types as originally observed by Gradie and Tedesco (1982). Further works by Mothé-Diniz et al. (2003), Carvano et al. (2010) and DeMeo and Carry (2014) demonstrate that the level of mixing increases for smaller asteroid sizes.

This main belt stratification makes intuitive sense in terms of a general view that proto-planetary disks should have temperatures decreasing with increasing distance from the central star. In fact, ordinary chondrites are less abundant in organics and water than carbonaceous chondrites and therefore are more likely to have formed in a warmer part of the disk. The small water content in ordinary chondrites, well below the solar proportion, suggests that these bodies accreted closer to the Sun than the snowline. The fact that some water is nevertheless present is not in contradiction with this statement. A small amount of water could have been accreted by collisions with primitive bodies scattered or drifting into the inner part of the disk. At the opposite extreme, the CI meteorites show no chemical fractionation relative to the solar composition, except H, C, N, O and all noble gases, suggesting that they formed in a region of the disk where the temperature was low enough to allow the condensation of most elements.

As shown in Fig. 3.6, however, asteroids of different taxonomic types are partially mixed in orbital semi major axis, which smears the trend relating physical properties to heliocentric distance. This pattern has not been interpreted to mean that asteroids of intermediate physical properties reside between major categories. Instead, it is likely due to the actual existence of asteroids of different taxonomic types in those semimajor axis zones. It is possible that some mixing could be a function of how the thermal and compositional properties of the disk evolved with time. Given that no systematic differences in accretion ages has yet been found among the main group of chondrites (i.e., most chondrules formed between 1-3 Myr after CAIs; Villeneuve et al., 2009), however, it seems more likely that some mechanism, possibly the same that excited the asteroids' orbital eccentricities and inclinations, also led to some modest stirring of their original semimajor axes.

3.5 Modeling work matched to constraints

In this section we review the most established models for the evolution of the asteroid belt, but we also discuss the potential of new ideas and the issues that remain to be explored. We break this discussion in two parts: first we address Scenario 2, the case where the asteroid belt supposedly formed with a low mass, then Scenario 1, the case where the belt was initially massive and had to be dynamically depleted fast enough to experience limited collisional activity.

3.5.1 From a low-mass asteroid belt

A primary challenge for models that hope to form a low-mass asteroid belt (Scenario 2) is to reproduce its observed orbital excitation.

If there is little mass in the belt, the self-stirring of asteroid orbits is extremely inefficient. According to Wetherill and Stewart (1993b) and Stewart and Ida (2000), the equations for the self-excitation of the mean eccentricity and inclination of of a population of planetesimals of individual mass M are:

$$\frac{\mathrm{d}e^2}{\mathrm{d}t} = \frac{C}{4} M e^2 (J_r(\beta) + 4J_\theta(\beta)) \, , \tag{3}$$

$$\frac{\mathrm{d}I^2}{\mathrm{d}t} = \frac{C}{4\beta^2} M I^2 J_z(\beta) \, . \tag{4}$$

Here $\beta = (I^2/e^2)^{1/2}$ and the functions J_r, J_θ and J_z are given in appendix A3 of Kenyon and Luu (1999). the coefficient C in eq. (3), (4) is given by Wetherill and Stewart (1993b):

$$C = \frac{16 G^2 \rho}{V_K^3 (2e^2)^{3/2}} (\log \Lambda + 0.55) \, , \tag{5}$$

where G is the gravitational constant, V_K is the Kepler velocity, Λ describes the minimum two-body deflection angle (detailed in Wetherill and Stewart, 1989) and ρ is the spatial density of particles j (this is the term where the total population of bodies of individual mass M intervenes).

Using these formulae, one can compute that an asteroid belt dominated by 5 Ceres-mass objects (thus a belt only moderately more massive than the current one) leads to limited self-stirring, such that the mean eccentricity of an initially cold main belt population would only get to 0.025 in 4.5 Gy, while the mean inclination would only get to 0.6 deg. Comparable results can also be obtained by tracking the effect of Ceres' perturbations on asteroids using direct numerical integration (e.g., Carruba et al. 2003). In order to stir up the main belt to a mean eccentricity of 0.1, the number of Ceres-mass bodies would need to be the order of $\sim 1,000$, or it would need to contain a few larger bodies for a more limited time. Either way, the belt would need to be much more massive than observed. In contrast, the low total mass of the cold Kuiper belt population is very much in line with this argument.

A possible way out of this is to argue that the low-mass asteroid belt was dynamically excited by external processes. Two processes have been proposed so far in the literature: (i) the sweeping of secular resonances through the belt during the dispersal of the gas from the protoplanetary disk (Heppenheimer, 1980; Ward, 1981; Lecar and Franklin, 1997; Nagasawa et al., 2000, 2001, 2002) and (ii) gravitational interactions between asteroids and massive bodies from the terrestrial planet and/or gas giant regions (e.g., Ip, 1987; Petit et al., 1999; Ward 2001; Levison et al. 2015b). As discussed in a review by Petit et al. (2002), the models tested at that time either had yet to produce a satisfactory eccentricity and inclination distribution or they had not yet been tested against all constraints.

For example, secular resonance sweeping from (i) does not reproduce the main belt's inclination distribution (O'Brien et al., 2007): apparently only a very slow gas-dispersal would be capable of substantially exciting asteroid inclinations, but at the cost of not producing a sufficient dispersion of final inclination values around the mean value (i.e. all asteroids have approximately the same inclination in the end, whereas the real ones have inclinations ranging from 0 to 30 degrees). Moreover, secular resonance sweeping does not produce radial mixing of asteroids of different taxonomic types; it instead preserves the initial semi-major axis distribution (e.g., O'Brien et al. 2007 and numerous references therein). A similar problem would likely be faced by invoking resonance excitation via the large embryos that presumably went on to form Jupiter's core (Ward 2001), though this has yet to be tested. (Note that O'Brien et al., (2007) assumed their giant planets were closer together and were on low eccentricity and low inclination orbits, which this explains some of the differences between their work and previous work; see Morbidelli et al. (2015) for further discussion).

The scattering of massive bodies from the giant planet region from (ii) would likely give the belt an uneven excitation and depletion distribution, with the outer part left notably more excited than the inner part because the outer part is more accessible to Jupiter-scattered bodies (Petit et al. 1999). The effect of planetary embryos scattered out of the terrestrial planet region has yet to be tested, though by definition, that model would also need to reproduce the low mass of Mars and the asteroid belt. Thus, models advocating an initially low-mass asteroid belt can be validated only if they can successfully address the orbital excitation and radial mixing constraints.

3.5.2 From a massive asteroid belt

If the belt was originally massive, the first constraint that needs to be addressed is its mass depletion (Scenario 1). More than 99% of the initial mass needs to be removed and this removal has to be fast enough to avoid too much collisional evolution, as discussed above.

Three mechanisms have been proposed in the literature, which we describe below.

3.5.3 Migration of planetary embryos

Ogihara et al. (2015) showed that planetary embryos originally in the asteroid belt, if sufficiently massive, can migrate out of the belt and into the terrestrial planet region by tidal interactions with the gas disk (so-called Type-I migration; e.g. Tanaka et al., 2002). Thus, if the planetary embryos carry the vast majority of the initial mass of the main belt, the final main belt would be left severely mass-depleted.

A problem with this scenario is that the rapid migration of planetary embryos out of the main belt would not provide sufficient orbital excitation of the remaining asteroids (Ogihara, private communication). In essence, the embryos leave the belt so quickly by Type-1 migration that they do not have time to excite the bodies left behind that are unaffected by this type of migration. Moreover, gas would still be present in the disk, since it would be required to drive the migration of the embryos, and therefore the gas drag would have helped damp the eccentricities and inclinations of the small bodies. Thus, this model would requires an external excitation mechanism similar to those discussed in the case of an initially low-mass asteroid belt (Scenario 2).

3.5.4 Stirring from a population of resident embryos

In this model, originally proposed by Wetherill (1992), when gas was removed from the Solar System, the protoplanetary disk interior to Jupiter consisted of a bi-modal population of planetesimals and planetary embryos, the latter with masses comparable to those of the Moon or Mars. Numerical simulations (Chambers and Wetherill, 1998; Petit et al., 2001; O'Brien et al., 2006, 2007) show that, under the effect of the mutual perturbations among the embryos and the resonant perturbations from Jupiter, embryos may leave the asteroid belt region, whereas they collide with each other, are scattered out of the solar system by Jupiter, or go on to build the terrestrial planets inside of 2 AU. While they are still crossing the asteroid belt, the embryos also excite and eject most of the original resident planetesimals. Only a minority of the planetesimals (and often no embryos) remain in the belt at the end of the terrestrial planets formation process, which explains the mass depletion of the current asteroid population. The eccentricities and inclinations of the surviving asteroids are excited and randomized, and the remaining asteroids have generally been scattered somewhat relative to their original semimajor axes, reproducing the observed mixing of taxonomic types.

This model is therefore quite successful in reproducing all the main properties of the asteroid belt: mass depletion, excitation, and radial mixing. Its main limitations are that (i) this model is not fully consistent with terrestrial planet formation because it tends to produce planets at the location of Mars which are too massive (Raymond et al., 2009); (ii) in several simulations embryos remain in the inner asteroid belt, particularly if Jupiter and Saturn were at the time on quasi-circular and co-planar orbits (Raymond et al., 2009; Izidoro et al., 2015); (iii) the surviving asteroids suffer intense collisional evolution because the dynamical depletion timescale is relatively slow (i.e., many tens of My). Thus, for the integrated collisional activity of asteroids to remain within the ~ 10 Gy constraint described above, the initial mass in planetesimals in the asteroid belt region probably had to have been no larger than 200 times the current asteroid belt mass, or less than one Mars mass (Bottke et al., 2005b). This implies that, if the belt originally had a mass of the order of an Earth mass, more than 90% of its primordial mass had to be in planetary embryos.

The most serious problem here is with constraint (i), namely that embryos in the asteroid belt often lead to a Mars that is much larger than that observed. If the small-Mars problem was solved by other mechanisms, however, Wetherill's model could be considered a valid possibility for dynamically sculpting the primordial asteroid belt.

3.5.5 Migration of Jupiter through the asteroid belt

This model, originally proposed in Walsh et al. (2011) and known colloquially as the 'Grand Tack' scenario, is built on results from hydrodynamics simulations that show that Jupiter migrates towards the Sun if it is alone in the gas-disk, while it migrates outward if paired with Saturn (Masset and Snellgrove, 2001; Morbidelli and Crida, 2007; Pierens and Nelson, 2008; Pierens and Raymond, 2011; D'angelo and Marzari, 2012). Thus, the Grant Tack postulates that Jupiter formed first and migrated inward. As long as Jupiter was basically alone, with early Saturn too small to substantially influence Jupiter's dynamics, Jupiter migrated inwards from its initial position (poorly constrained but estimated at ∼3.5 AU) down to 1.5 AU. Then, when Saturn approached its current mass and migrated inward to an orbit close to that of Jupiter, Jupiter reversed migration direction (aka it "tacked", hence the name of the model). This allowed the pair of planets to both move outwards together. This migration continued until all gas was removed from the disk, which the model assumed took place when Jupiter reached a distance of ∼5.5 AU.

The Grand Tack model assumes that Jupiter formed just outside the snowline at 3.5 AU. The planetesimals originally inside its initial orbit were assumed to be predominantly S-complex, some which may be oridinary chondrites, others which may be highly metamorphosed bodies (e.g., enstatite chondrite, Earth precursors, etc.). During its inward migration, Jupiter penetrates into the disk of these planetesimals (whose distribution is sketched as a dashed area in Fig. 3.7). In doing so, most planetesimals are captured in mean motion resonances with Jupiter and are pushed inwards. However, some 10% of the planetesimals are kicked outwards by an encounter with Jupiter, reaching orbits located beyond Saturn, which collectively have an orbital (a,e) distribution that is typical of a scattered disk (i.e. with mean eccentricity increasing with semimajor axis). In semimajor axis range, this scattered disk overlaps with the inner part of the disk of primitive bodies (whose distribution is sketched as a dotted area in Fig. 3.7), which are initially on circular orbits beyond the orbit of Saturn.

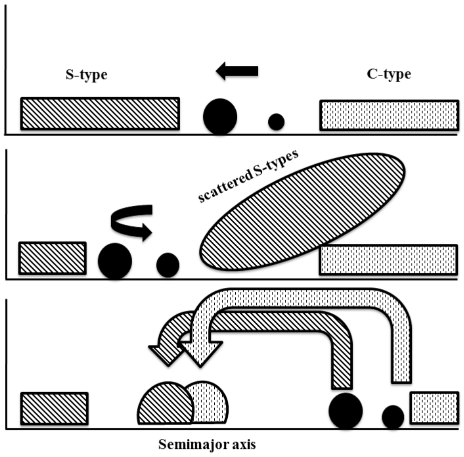

Fig. 3.7.— A scheme showing the Grand Tack evolution of Jupiter and Saturn and its effects on the asteroid belt. The three panels show three evolutionary states, in temporal sequence. First the planet migrate inwards then, when Saturn reaches its current mass, they move outwards. The dashed and dotted areas schematize the (a,e) distributions of S-complex and C-complex asteroids respectively. The dashed and dotted arrows in the lower panel illustrate the injection of scattered S-complex and C-complex asteroids into the asteroid belt during the final phase of outward migration of the planets.

These bodies, being formed beyond the snowline, should be rich in water ice and other volatile elements, and therefore the model associates them with C-complex asteroids. After Jupiter reaches $\sim$1.5 AU (this value is constrained by the requirement to form a small Mars and a big Earth; Walsh et al., 2011; Jacobson et al., 2014; Jacobson and Morbidelli, 2014), Saturn's inward migration changes the structure of how the planets interact gravitationally with the disk. This causes Jupiter to reverse its migration direction and move outward. During this time, the Jupiter and Saturn encounter the scattered S-complex disk, and then also the primitive C-complex disk. Some of the bodies in both populations are kicked inwards, where they reach the asteroid belt region and are implanted there as Jupiter moves out of it.

The migration of Jupiter through the asteroid belt region leaves the final asteroid belt highly depleted in mass. In fact, the probability that a S-complex body is scattered back into the asteroid belt is of the order of a few times 10^{-3} (Walsh et al., 2011). This dynamical depletion occurs quickly (~ 0.1 My) and very early in the history of the Solar System. By definition, it has to occur before the complete removal of the gas disk, which is thought to last 3-4 My after the formation of first solids (i.e., the CAIs). Collisional evolution during this time among asteroids is intense but brief, and the surviving bodies in the main belt should fulfill the 10 Gy constraint on the integrated collisional activity of asteroids, as described above (Morbidelli et al., 2015b). Modeling this scenario with collision evolution included and checking it again all main belt constraints, however, has yet to be attempted.

The final orbital eccentricity and inclination distributions of the asteroids in the belt appear to be excited and randomized (Walsh et al., 2011). A potential concern is that the final eccentricity distribution is skewed towards the high-eccentricity boundary of the asteroid belt, whereas the observed one peaks around ~ 0.1. If one accounts for the subsequent evolution of the asteroids, however, it may be possible to lower these values. Possible mechanisms include a later phase of dynamical instability among the giant planets (e.g., the Nice model; Morbidelli et al., 2010; 2015). As the giant planets undergo an orbital reshuffling, some asteroids in the main belt may undergo chaotic diffusion; some bodies will evolve to lower eccentricities while others reach planet-crossing orbits and escape into the terrestrial planet region. This process is the subject of ongoing work, though preliminary work suggests the model asteroid distribution can approach the observed one (Deienno and Gomes, personal communication).

The mixing of taxonomic types is achieved in the Grand Tack model by injecting many C-complex asteroids into the asteroid belt during Jupiter's outward migration phase. The final distribution (Walsh et al., 2011) shows that S-complex asteroids dominate the inner belt and C-complex asteroids dominate the outer belt, but both populations overlap over the entire asteroid belt semi major axis range. The inclusion of additional C-complex bodies during a late dynamical instability of the giant planets is also a possibility (Levison et al. 2009).

3.6 Summary

The asteroid belt has been a key witness to many of the major collisional and dynamical events that have taken place in the inner solar system over its history. This means our planetesimal and planet formation models should not be considered complete until they can satisfy the numerous constraints provided by this population. While considerable progress has been made over the last several decades in interpreting how the main belt reached its current state, fundamental issues still await resolution. For example, as described here, there is currently a debate on whether the primordial main belt was initially massive or whether it has always been close to its present-day mass. In addition, the nature of the dynamical processes needed for the asteroid belt to achieve its current dynamically excited and semi-mixed state – in terms of sizable S- and C-complex asteroids – are still being studied. Our interpretation of the collisional evolution of the main belt also depends on constraints like the number and sizes of asteroid families that may become increasingly incomplete as we go further back in time. Fortunately, or unfortunately, depending on your point of view, all of these questions are intertwined, so an advance on one problem may allow us to more readily address the others.

Additional headway on these issues may also come from the application of new constraints. For example, many meteorites came from large main belt asteroids via a collisional cascade (Bottke et al. 2005c). Studies of these tiny asteroid samples provide us with a treasure trove of data that be used to help us understand both planetesimal formation and evolution (e.g., how these bodies were affected by early bombardment; Marchi et al. 2013; Bottke et al. 2015a). The issue is placing these data into the correct solar system context, which is a job for our collisional and dynamical evolution models. Progress could also come from new missions to large primordial asteroids, some which may possess critical clues that bear on the issues discussed here. Our visits to Vesta, Ceres, and Lutetia via NASA's Dawn and ESA's Rosetta missions have only begun to whet our appetite. There is also much that can still be accomplished with ground- and space-based observation campaigns. Some key examples are the asteroid color data provided by the ground-based Sloan Digital Sky Survey (Parker et al. 2008) and the asteroid albedo data provided by NASA's space-based infrared telescopic survey WISE (e.g., Masiero et al. 2011).

Perhaps the most important issue of all, however, is to consider the asteroid belt as part of our entire system of worlds that stretches from Mercury to the Oort cloud. Major dynamical events often affect multiple worlds, so the most powerful and insightful models are those that can match all of the available constraints, not just those for individual worlds or a limited number of asteroids. Only then will we be able to say with some confidence that we have answers that are likely to be robust.

Acknowledgments We thank our reviewers Dave O'Brien and Erik Asphaug for their helpful and constructive comments. Research funds for William Bottke were provided by NASA's Solar System Evolution Research Virtual Institute (SSERVI) as part of the Institute for the Science of Exploration Targets (ISET) at the Southwest Research Institute (NASA grant number NNA14AB03A).

REFERENCES

Anders, E., 1965. Fragmentation history of asteroids. Icarus 4, 399-408.

Asphaug, E., G. Collins, and M. Jutzi, 2015. Global-Scale Impacts. In Asteroids IV (P. Michel, F. DeMeo, and W. F. Bottke, Eds). U. Arizona Press, in press.

Benz, W., and E. Asphaug, 1999. Catastrophic disruptions revisited. Icarus 142, 5-20.

Binzel, R. P., Bus, S. J., Burbine, T. H., Sunshine, J. M., 1996. Spectral properties of near-Earth asteroids: Evidence for sources of ordinary chondrite meteorites. Science 273, 946-948.

Bottke, W. F., Durda, D. D., Nesvorný, D., et al., 2005a. The and fossilized size distribution of the main asteroid belt. Icarus 175, 111-140.

Bottke, W. F., Durda, D. D., Nesvorný, D., et al., 2005b. Linking the collisional history of the main asteroid belt to its dynamical excitation and depletion. Icarus 179, 63-94.

Bottke, W. F., Durda, D. D., Nesvorný, D., et al., 2005c, The origin and evolution of stony meteorites, in *Dynamics of Populations of Planetary Systems*, ed. Z. Knežević and A. Milani (Cambridge University Press, Cambridge), 357–374.

Bottke, W. F., Nesvorný, D., Grimm, R. E., et al., 2006. Iron meteorites as remnants of planetesimals formed in the terrestrial planet region. Nature 439, 821-824.

Bottke, W. F., Levison, H. F., Nesvorný, D., Dones, L., 2007. Can planetesimals left over from terrestrial planet formation produce the lunar Late Heavy Bombardment?. Icarus 190, 203-223.

Bottke, W. F., Vokrouhlický, D., Minton, D., et al., 2012. An Archaean heavy bombardment from a destabilized extension of the asteroid belt. Nature 485, 78-81.

Bottke, W. F., Vokrouhlický, D., Marchi, S., et al., 2015a. Dating the Moon-forming impact event with asteroidal meteorites. Science 348, 321-323.

Bottke, W. F., Broz, M., O'Brien, D. P., et al., 2015b. The collisional evolution of the asteroid belt. In Asteroids IV (P. Michel, F. DeMeo, and W. F. Bottke, Eds). U. Arizona Press, in press.

Brown, P. G., Assink, J. D., Astiz, L., et al., 2013. A 500-kiloton airburst over Chelyabinsk and an enhanced hazard from small impactors. Nature 503, 238-241.

Brož, M., Morbidelli, A., Bottke, W.F., et al., 2013. Constraining the cometary flux through the asteroid belt during the late heavy bombardment, Astron. Astrophys. 551, A117.

Burbine, T. H., McCoy, T. J., Meibom, et al., 2002. Meteoritic parent bodies: Their number and identification. In Asteroids III (W. F. Bottke et al., eds). U. Arizona Press, Tucson, 653-667.

Buczkowski, D. L., Wyrick, D. Y., Iyer, K. A., et al., 2012. Large-scale troughs on Vesta: A signature of planetary tectonics. Geophysical Research Letters 39, L18205.

Bus, S. J., Binzel, R. P., 2002. Phase II of the small main-belt asteroid spectroscopic survey. A feature-based taxonomy. Icarus 158, 146-177.

Carvano, J. M., Hasselmann, P. H., Lazzaro, D., Mothé-Diniz, T., 2010. SDSS-based taxonomic classification and orbital distribution of main belt asteroids. Astronomy and Astrophysics 510, A43.

Chambers, J. E. and Wetherill, G. W., 1998. Making the terrestrial Planets: N-body integrations of planetary embryos in three dimensions. Icarus 136, 304-327.

Cibulková, H., Brož, M., Benavidez, P. G., 2014. A six-part collisional model of the main asteroid belt. Icarus 241, 358-372.

Connelly, J. N., Bizzarro, M., Krot, A. N., et al., 2012. The absolute chronology and thermal processing of solids in the solar protoplanetary disk. Science 338, 651-654.

Cuzzi, J. N., Hogan, R. C., Bottke, W. F., 2010. Towards initial mass functions for asteroids and Kuiper Belt Objects. Icarus 208, 518-538.

D'Angelo, G., and Marzari, F., 2012. Outward migration of Jupiter and Saturn in evolved gaseous disks. Astrophys. J. 757, 50.

Davison, T. M., O'Brien, D. P., Ciesla, F., et al., 2013. The early impact histories of meteorite parent bodies. Meteoritics and Planetary Science 48, 1894-1918.

Davis, D. R., Durda, D. D., Marzari, F., et al., 2002. Collisional evolution of small body populations. In Asteroids III (W.F. Bottke, et al., eds.), U. Arizona Press, Tucson, 545-558.

DeMeo, F. E., Binzel, R. P., Slivan, S. M., Bus, S. J., 2009. An extension of the Bus asteroid taxonomy into the near-infrared. Icarus 202, 160-180.

DeMeo, F. E., Carry, B., 2014. Solar System evolution from compositional mapping of the asteroid belt. Nature 505, 629-634.

Durda, D. D., Bottke, W. F., Enke, B. L., et al., 2004. The formation of asteroid satellites in large impacts: results from numerical simulations. Icarus 170, 243-257.

Durda, D.D., et al., 2007. Size-frequency distributions of fragments from SPH/N-body simulations of asteroid impacts: Comparison with observed asteroid families, *Icarus* **186**, 498–516.

Eugster, O., 2003. Cosmic-ray exposure ages of meteorites and lunar rocks and their significance. Chemie der Erde 63, 3-30.

Fowler, J.W. and Chillemi, J. R., 1992. IRAS asteroid data processing. In The IRAS Minor Planet Survey (E.F. Tedesco, ed) Tech. Report PL-TR-92-2049., Phillips Laboratory, Hanscom Air Force Base, Massachusetts, 17-43.

Fraser, W. C., Brown, M. E., Morbidelli, A., et al., 2014. The absolute magnitude distribution of Kuiper belt objects. Astrophys. J. 782, 100.

Gradie, J., Tedesco, E., 1982. Compositional structure of the asteroid belt. Science 216, 1405-1407.

Güttler, C., Blum, J., Zsom, A., et al., 2009. The first phase of protoplanetary dust growth: The bouncing barrier. Geochimica et Cosmochimica Acta Supplement 73, 482.

Harris, A. W., and D'Abramo, G., 2015. The population of near-Earth asteroids. Icarus 257, 302-312.

Harris, A. W., Boslough, M. Chapman, C. R., et al., 2015. Asteroid impacts and modern civilization: Can we prevent a catastrophe? In Asteroids IV (P. Michel, et al., eds). U. Arizona Press, Tucson, in press.

Hartmann, W. K., and Hartmann, A. C. 1968. Asteroid collisions and evolution of asteroidal mass distribution and meteoritic flux. Icarus 8, 361-381.

Hayashi, C., 1981. Structure of the solar nebula, growth and decay of magnetic fields and effects of magnetic and turbulent viscosities on the nebula. Progress of Theoretical Physics Supplement 70, 35-53.

Heppenheimer, T. A., 1980. Secular resonances and the origin of eccentricities of Mars and the asteroids. Icarus, 41, 76-88.

Hiroi, T., Zolensky, M. E., Pieters, C. M., 2001. The Tagish Lake meteorite: A possible sample from a D-Type Asteroid. Science 293, 2234-2236.

Ip, W.-H., 1987. Gravitational stirring of the asteroid belt by

Jupiter zone bodies. Beitraegezur Geophysik 96, 44-51.

Ivanov, B. A., Neukum, G., Bottke, W. F., and Hartmann, W. K. 2002. The comparison of size-frequency distributions of impact craters and asteroids and the planetary cratering rate. In Asteroids III, (W. F. Bottke, et al., eds) U. Arizona Press, Tucson, 89-101.

Izidoro, A., Raymond, S. N., Morbidelli, A., Winter, O. C., 2015. Terrestrial planet formation constrained by Mars and the structure of the asteroid belt. Monthly Notices of the Royal Astronomical Society 453, 3619-3634.

Jacobson, S. A., Morbidelli, A., Raymond, S. N., et al., 2014. Highly siderophile elements in Earth's mantle as a clock for the Moon-forming impact. Nature 508, 84-87.

Jacobson, S. A., and Morbidelli, A., 2014. Lunar and terrestrial planet formation in the Grand Tack scenario. Royal Society of London Philosophical Transactions Series A 372, 174.

Jaumann, R., Williams, D. A., Buczkowski, D. L., et al., 2012. Vesta's shape and morphology. Science 336, 687-690.

Jedicke, R., Larsen, J., and Spahr, T, 2002. Observational selection effects in asteroid surveys. In Asteroids III (W. F. Bottke, et al, eds.) U. Arizona Press, Tucson, 71-88.

Johansen, A., Henning, T., Klahr, H., 2006. Dust sedimentation and self-sustained Kelvin-Helmholtz turbulence in protoplanetary disk midplanes. Astrophys. J. 643, 1219-1232.

Johansen, A., Oishi, J. S., Mac Low, M.-M., et al., 2007. Rapid planetesimal formation in turbulent circumstellar disks. Nature 448, 1022-1025.

Johansen, A., Mac Low, M.-M., Lacerda, P., Bizzarro, M., 2015. Growth of asteroids, planetary embryos, and Kuiper belt objects by chondrule accretion. Science Advances 1, 1500109.

Lecar, M. and Franklin, F., 1997. The solar nebula, secular resonances, gas drag, and the asteroid belt. Icarus 129, 134-146.

Kenyon, S. J., and Luu, J. X., 1999. Accretion in the early outer solar system. Astrophys. J. 526, 465-470.

Levison, H. F., Bottke, W. F., Gounelle, M., et al., 2009. Contamination of the asteroid belt by primordial trans-Neptunian objects. Nature 460, 364-366.

Levison, H. F., Kretke, K. A., Duncan, M. J. 2015. Growing the gas-giant planets by the gradual accumulation of pebbles. Nature 524, 322-324.

Levison, H. F., Kretke, K. A., Walsh, K., Bottke, W. F., 2015b. Growing the terrestrial planets from the gradual accumulation of submeter-sized objects. Proc. Nat. Acad. Sciences, in press.

Kretke, K. A., Lin, D. N. C., 2007. Grain retention and formation of planetesimals near the snow line in MRI-driven turbulent protoplanetary disks. Astrophys. J. 664, L55-L58.

Kuiper, G. P., Fugita, Y. F., Gehrels, T., et al., 1958. Survey of asteroids. Astrophysical Journal Supplement Series 3, 289.

Mainzer, A., Trilling, D., and Usui, F., 2015. Space-based infrared studies of asteroids. In Asteroids IV (P. Michel, et al., eds) U. Arizona Press, Tucson, in press.

Marchi, S., McSween, H. Y., O'Brien, D. P., et al., 2012. The violent collisional history of asteroid 4 Vesta. Science 336, 690-693.

Marchi, S., Bottke, W. F., Cohen, B. A., 2013. High-velocity collisions from the lunar cataclysm recorded in asteroidal meteorites. Nature Geoscience 6, 303-307.

Masiero, J., Mainzer, A. K., Grav, T., 2011. Main belt asteroids with WISE/NEOWISE I: Preliminary albedos and diameters. Astrophys. J. 741, 68.

Masiero, J. R., DeMeo, F., Kasuga, T. and Parker, A. H. 2015. Asteroid family physical properties. In Asteroids IV (P. Michel, et al., eds). U. Arizona Press, Tucson, in press.

Masset, F., and Snellgrove, M., 2001. Reversing type II migration: resonance trapping of a lighter giant protoplanet. Monthly Notices of the Royal Astronomical Society 320, L55-L59.

McEwen, A. S., Moore, J. M., Shoemaker, E. M., 1997. The Phanerozoic impact cratering rate: Evidence from the farside of the Moon. Journal of Geophysical Research 102, 9231-9242.

Morbidelli, A. and Gladman, B. 1998. Orbital and temporal distributions of meteorites originating in the asteroid belt. Meteoritics and Planetary Science 33, 999-1016

Morbidelli, A. and Vokrouhlický, D., 2003. The Yarkovsky-driven origin of near-Earth asteroids. Icarus 163, 120-134.

Morbidelli, A., and Crida, A., 2007. The dynamics of Jupiter and Saturn in the gaseous protoplanetary disk. Icarus 191, 158-171.

Morbidelli, A., Brasser, R., Gomes, R., et al., 2010. Evidence from the asteroid belt for a violent past evolution of Jupiter's orbit. Astron. J. 140, 1391-1401.

Morbidelli, A., Lambrechts, M., Jacobson, S., Bitsch, B., 2015. The great dichotomy of the Solar System: small terrestrial embryos and massive giant planet cores. Icarus, in press.

Morbidelli, A., Walsh, K. J., O'Brien, D. P., et al., 2015b. Dynamical evolution of the asteroid belt. In Asteroids IV (P. Michel, et al., eds), U. Arizona Press, Tucson, in press.

Mothé-Diniz, T., Carvano, J. M. á., Lazzaro, D., 2003. Distribution of taxonomic classes in the main belt of asteroids. Icarus 162, 10-21.

Nagasawa, M., Tanaka, H., and Ida, S., 2000. Orbital evolution of asteroids during depletion of the solar nebula. Astron. J., 119, 1480-1497.

Nagasawa, M., Ida, S., and Tanaka, H., 2001. Origin of high orbital eccentricity and inclination of asteroids. Earth, Planets, and Space, 53, 1085-1091.

Nagasawa, M., Ida, S., and Tanaka, H., 2002. Excitation of orbital inclinations of asteroids during depletion of a protoplanetary disk: Dependence on the disk configuration. Icarus, 159, 322-327.

Nesvorný, D., Youdin, A. N., Richardson, D. C., 2010. Formation of kuiper belt binaries by gravitational collapse. Astron. J. 140, 785-793.

Nesvorný, D., Vokrouhlický, D., Bottke, W. F., et al., 2011. Observed binary fraction sets limits on the extent of collisional grinding in the Kuiper belt. Astron. J. 141, 159.

Nesvorný, D., Brož, M., and Carruba, V. 2015. Identification and dynamical properties of asteroid families. In Asteroids IV (P. Michel, et al., eds). U. Arizona Press, Tucson, in press.

Noll, K. S., Parker, A. H., Grundy, W. M., 2014. All bright cold classical KBOs are binary. AAS/Division for Planetary Sciences Meeting Abstracts 46, 507.05.

O'Brien, D. P., Morbidelli, A., and Levison, H. F., 2006. Terrestrial planet formation with strong dynamical friction. Icarus, 184, 39-58.

O'Brien, D. P., Morbidelli, A., and Bottke, W. F., 2007. The primordial excitation and clearing of the asteroid belt-Revisited. Icarus, 191, 434-452.

Ogihara, M., Kobayashi, H., Inutsuka, S.-i., Suzuki, T. K., 2015. Formation of terrestrial planets in disks evolving via disk winds and implications for the origin of the solar system's terrestrial planets. Astron. Astrophys. in press.

Parker, A., Ivezić, Ž., Jurić, M., et al., 2008, The size distributions of asteroid families in the SDSS Moving Object Catalog 4, Icarus 198, 138-155.

Parker, A. H., Kavelaars, J. J., Petit, J.-M., et al., 2011. Characterization of seven ultra-wide trans-Neptunian binaries. Astrophys. J. 743, 1.

Petit, J., Morbidelli, A., and Valsecchi, G. B., 1999. Large scattered planetesimals and the excitation of the small body belts. Icarus 141, 367-387.

Petit, J., Morbidelli, A., and Chambers, J., 2001. The primordial excitation and clearing of the asteroid belt. Icarus 153, 338-347.

Petit, J., Chambers, J., Franklin, F., and Nagasawa, M., 2002. Primordial excitation and depletion of the main belt. In Asteroids III (W. F. Bottke, et al., eds) U. Arizona Press, Tucson, 711-738.

Petit, J.-M., Mousis, O., 2004. KBO binaries: How numerous were they? Icarus 168, 409-419.

Pierens, A., Nelson, R. P., 2008. Constraints on resonant-trapping for two planets embedded in a protoplanetary disc. Astronomy and Astrophysics 482, 333-340.

Pierens, A., Raymond, S. N., 2011. Two phase, inward-then-outward migration of Jupiter and Saturn in the gaseous solar nebula. Astronomy and Astrophysics 533, A131.

Raymond, S. N., O'Brien, D. P., Morbidelli, A., Kaib, N. A., 2009. Building the terrestrial planets: Constrained accretion in the inner Solar System. Icarus 203, 644-662.

Rabinowitz, D. Helin, L. E., Lawrence, K., and Pravdo., S., 2000. A Reduced estimate of the number of kilometre-sized near-Earth asteroids. Nature 403, 165-166

Russell, C. T., H. Y. McSween, R. Jaumann, and C. A. Raymond. 2015. The Dawn mission to Vesta and Ceres. In Asteroids IV (P. Michel, et al., eds). U. Arizona Press, Tucson, in press.

Schenk, P., O'Brien, D. P., Marchi, D., et al., 2012. The geologically recent giant impact basins at Vesta's south pole. Science 336, 694-697.

Scott, E. R. D., Keil, K., Goldstein, J. I., et al., 2015. Early impact history and dynamical origin of differentiated meteorites and asteroids. In Asteroids IV (P. Michel, et al., eds). U. Arizona Press, Tucson, in press.

Stewart, G. R., and Ida, S. 2000. Velocity evolution of planetesimals: Unified analytical formulas and comparisons with N-body simulations. Icarus 143, 28-44.

Stöffler, D., and Ryder, G., 2001. Stratigraphy and isotope ages of lunar geologic units: Chronological standard for the inner solar system. Space Science Reviews 96, 9-54.

Stokes, G. H., Yeomans, D. K., Bottke, W. F., et al., 2003, Report of the Near-Earth Object Science Definition Team: A Study to Determine the Feasibility of Extending the Search for Near-Earth Objects to Smaller Limiting Diameters. NASA-OSS-Solar System Exploration Division.

Tanaka, H., Takeuchi, T., Ward, W. R., 2002. Three-dimensional interaction between a Planet and an isothermal gaseous disk. I. Corotation and Lindblad torques and planet migration. Astrophys. J. 565, 1257-1274.

Tholen, D. J., 1984. Asteroid taxonomy from cluster analysis of Photometry. Ph.D. Thesis.

Turrini, D., Magni, G., Coradini, A., 2011. Probing the history of Solar system through the cratering records on Vesta and Ceres. Monthly Notices of the Royal Astronomical Society 413, 2439-2466.

Turrini, D., Coradini, A., Magni, G., 2012. Jovian early bombardment: Planetesimal erosion in the inner asteroid belt. Astrophys. J. 750, 8.

Vokrouhlický, D., Broz, M., Bottke, W. F., et al., 2006. Yarkovsky/YORP chronology of asteroid families. Icarus 182, 118-142.

Vokrouhlický, D., Bottke, W. F., Chesley, S. R., et al., 2015. The Yarkovsky and YORP effects. In Asteroids IV (P. Michel, et al., eds). U. Arizona Press, Tucson, in press.

Villeneuve, J., Chaussidon, M., Libourel, G., 2009. Homogeneous distribution of ^{26}Al in the Solar System from the Mg isotopic composition of chondrules. Science 325, 985.

Walsh, K. J., Morbidelli, A., Raymond, S. N., et al., 2011. A low mass for Mars from Jupiter's early gas-driven migration. Nature 475, 206-209.

Ward, W. R., 1981. Solar nebula dispersal and the stability of the planetary system. I - Scanning secular resonance theory. Icarus, 47, 234-264.

Weidenschilling, S. J., 1977. The distribution of mass in the planetary system and solar nebula. Astrophysics and Space Science 51, 153-158.

Weidenschilling, S. J., 1977. Aerodynamics of solid bodies in the solar nebula. Monthly Notices of the Royal Astronomical Society 180, 57-70.

Weidenschilling, S. J., Spaute, D., Davis, D. R., et al., 1997. Accretional evolution of a planetesimal swarm. Icarus 128, 429-455.

Weidenschilling, S. J., 2011. Initial sizes of planetesimals and accretion of the asteroids. Icarus 214, 671-684.

Wetherill, G. W., 1989. Origin of the asteroid belt. Asteroids II 661-680.

Wetherill, G. W., Stewart, G. R., 1989. Accumulation of a swarm of small planetesimals. Icarus 77, 330-357.

Wetherill, G. W., 1992. An alternative model for the formation of the asteroids. Icarus 100, 307-325.

Wetherill, G. W., Stewart, G. R., 1993. Formation of planetary embryos - Effects of fragmentation, low relative velocity, and independent variation of eccentricity and inclination. Icarus 106, 190-209.

Wilhelms, D. E., Oberbeck, V. R., Aggarwal, H. R., 1978. Size-frequency distributions of primary and secondary lunar impact craters. Lunar and Planetary Science Conference 9, 1256-1258.

Wilhelms, D.E., 1987. The Geologic History of the Moon. U.S. Geol. Surv. Prof. Pap. 1348, 302 pp.

Youdin, A. N., and Goodman, J., 2005. Streaming instabilities in protoplanetary disks. Astrophys. J. 620, 459-469.

Youdin, A., and Johansen, A., 2007. Protoplanetary disk turbulence driven by the streaming instability: linear evolution and numerical methods. Astrophys. J. 662, 613-626.