1. INTRODUCTION

The asteroid belt helps us in reconstructing the origin and the evolution of the solar system, probably better than the planets themselves. This is because the asteroid belt provides several key constraints that can be used to effectively guide the development, calibration, and validation of evolutionary models. Compared to other small-body populations, such as the Kuiper belt or Oort cloud, the constraints provided by the asteroid belt are probably more stringent, due to the fact that the number and the properties of the asteroids are better known, thanks to groundbased observations, space missions, and meteorite analysis.

The structure of this review chapter is therefore as follows. We start by reviewing in section 2 what the most important observational constraints on the asteroid belt are and what they suggest. Then, in section 3, we will review the main models proposed, from the oldest to the most recent, and from the earliest to the latest evolutionary phases they address. In section 4, we will discuss several implications for asteroid science from our current preferred view of the dynamical evolution of the asteroid belt.

The dynamical evolution of the asteroid belt has already been the object of a review chapter by Petit et al. (2002) in the Asteroids III book. This review therefore has an important overlap with that chapter. Nevertheless, both our...
observational knowledge of the asteroid belt and our theoretical understanding of solar system evolution have improved significantly since the early 2000s, providing an emerging view of a very dynamic early solar system, in which various episodes of planet migration played a fundamental role in sculpting the small-body reservoirs and displacing planetesimals far from their original birthplaces. Thus this chapter will present in greater details models proposed after 2002, focusing on their implications for asteroid science. Moreover, when reviewing models already presented in Petit et al., we will refer to numerical simulations of these models made after the publication of the Petit et al. chapter.

2. OBSERVATIONAL CONSTRAINTS ON THE PRIMORDIAL EVOLUTION OF THE ASTEROID BELT

The observational constraints most useful for reconstructing the formation and evolution of the asteroid belt are those related to large asteroids (larger than ~50–100 km in diameter). In fact, it has been argued that these asteroids are the most likely to be “pristine” in the sense that they were not generated in large numbers in collisional breakup events of larger parent bodies (Bottke et al., 2005a; see also the chapter by Bottke et al. in this volume), nor have they been affected by gas drag and other nongravitational forces (e.g., the Yarkovsky effect; see the chapter by Vokrouhlický et al. in this volume). Moreover, there is an emerging view that the first planetesimals were big, with a preferred diameter in the range mentioned above (Morbidelli et al., 2009; see also the chapter by Johansen et al. in this volume). Thus, throughout this chapter we will limit our discussion to the properties of large asteroids and refer to smaller asteroids only when explicitly mentioned.

A key major characteristic of the asteroid belt population is the orbital excitation, i.e., the fact that the eccentricities and inclinations of many asteroidal orbits are quite large (e.g., Petit et al., 2002). The median proper inclination of D > 100 km asteroids is 11° and the median proper eccentricity is 0.145. More importantly, the values of eccentricities and inclinations of the largest asteroids are considerably dispersed, with the former ranging between 0 and 0.30, while the latter ranges between 0° and 33° (see Fig. 1). It has been shown that asteroids of modest inclinations (i < 20°) fill the entire orbital space available for long-term dynamical stability, although some stable regions are more densely populated than others (Minton and Malhotra, 2009, 2011). The reader should be aware that, whatever the preferred formation mechanism (see the chapter by Johansen et al. in this volume), planetesimals are expected to have formed on circular and coplanar orbits. Thus, one or more dynamical excitation mechanism(s) within the primordial asteroid belt were needed to stir up eccentricities and inclinations to randomly dispersed values. Asteroid eccentricities and inclinations do not show a strong dependence on semimajor axis (Fig. 1).

A second fundamental 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; see also the chapter by Vernazza et al. in this volume). The central belt (2.5–3.2 AU) is dominated by C-complex asteroids, probably related to carbonaceous chondrites (Burbine et al., 2002; see also the chapters by DeMeo et al. and Rivkin et al. in this volume). 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 (see the chapter by Emery et al. in this volume). The C2 ungrouped meteorite “Tagish Lake” has been proposed to be a fragment of a D-type asteroid (Hiroi et al., 2001).

This stratification of the main belt makes intuitive sense in terms of a general view that protoplanetary 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
orders of magnitude compared to these expectations. As shown in Fig. 2, however, asteroids of different taxonomic types are partially mixed in orbital semimajor axis, which smears the trend relating physical properties to heliocentric distance. The mixing of taxonomic type should not be interpreted as the existence of asteroids of intermediate physical properties between those of adjacent types; instead, it is due to the coexistence of asteroids of different types with various relative proportions at each value of semimajor axis. Some mixing could come from the fact that the thermal and chemical compositional properties of the disk probably changed over time. However, given that no systematic differences in accretion ages is observed among the main group of chondrites (Villeneuve et al., 2009), it is more likely that some mechanism, possibly the same that excited the orbital eccentricities and inclinations, also changed somewhat in a random fashion the original semimajor axes of the bodies, causing the observed partial mixing.

The asteroid belt contains overall very little mass. From the direct determination of the masses of the largest asteroids and an estimate of the total mass of the ring of bodies, which cannot be individually “weighted,” based on the collective gravitational perturbations exerted on Mars, Krasinsky et al. (2002), Kuchynka and Folkner (2013), and Somenzi et al. (2010) concluded that the total mass contained in the asteroid belt is \(4.5 \times 10^{-4}\) Earth masses (\(M_\oplus\)). This value is very low compared to that estimated to have originally existed in the primordial asteroid belt region by interpolating the mass densities required to form the terrestrial planets and the core of Jupiter at both ends of the belt (Weidenschilling, 1977), which is on the order of 1 \(M_\oplus\) (within a factor of a few). Thus, the mass in the asteroid belt has potentially been depleted by 3 orders of magnitude compared to these expectations.

We can glean insights into how the primordial belt lost its mass by investigating what we know about its collisional evolution. The collisional history of asteroids is the subject of the chapter by Bottke et al. in this volume, but we report the highlights here that are needed for this discussion. In brief, using a number of constraints, Bottke et al. (2005a) concluded that the integrated collisional activity of the asteroid belt is equivalent to the one that would be produced at the current collisional rate over 8–10 G.y.

This result has several implications. First, it strongly suggests that the 3-orders-of-magnitude mass depletion could not come purely from collisional erosion; such intense comminution would violate numerous constraints. Second, it argues that the mass depletion of the asteroid belt occurred very early. This is because once the eccentricities and inclinations are excited to values comparable to the current ones, for a given body every million years spent in an asteroid belt 1000× more populated brings a number of collisions equivalent to that suffered in 1 G.y. within the current population. For this reason, the third implication is that the dynamical excitation and the mass-depletion event almost certainly coincided. This argues that the real dynamical excitation event was stronger than suggested by the current distribution of asteroid eccentricities. One way to reconcile a massive asteroid belt with this scenario is to assume that more than 99% of the asteroids had their orbits so excited that they left the asteroid belt forever (hence the mass depletion). This would make the eccentricities (and to a lesser extent the inclinations) we see today to be those defined by the lucky survivors, namely the bodies whose orbits were excited the least.

Using these constraints, we discuss in the next section the various models that have been proposed for the primordial sculpting of the asteroid belt.

3. MODELS OF PRIMORDIAL EVOLUTION OF THE ASTEROID BELT

3.1. Early Models

The first attempts to explain the primordial dynamical excitation of the asteroid belt were made by Heppenheimer (1980) and Ward (1981), who proposed that secular resonances swept through the asteroid belt region during the dissipation of gas in the protoplanetary disk. Secular resonances occur when the precession rate of the orbit of an asteroid is equal to one of the fundamental frequencies of precession of the orbits of the planets. There are two angles that characterize the orientation of an orbit in space, the longitude of perihelion (\(\Omega\)) and the longitude of the ascending node (\(\Omega\)), each of which can precess at different rates depending on the gravitational effects of the other planets and nebular gas (if present). The resonances that occur when the precession rates of the longitudes of perihelion of an asteroid (denoted by \(g\)) and of a planet are equal to each other excite the asteroid’s eccentricity. Similarly, the resonances occurring when the precession rates of the longitudes of node of an asteroid (denoted by \(s\)) and of a planet are equal to each other.

![Fig. 2. 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.](image-url)
excite the asteroid’s inclination. In the case of asteroids in the main belt, the planets’ precession frequencies that most influence their dynamics are those associated with the orbits of Jupiter and Saturn. These are called \( g_s \) and \( g_a \), for the longitude of perihelion precession (the former dominating in the precession of the perihelion of Jupiter, the latter in that of Saturn), and \( s_p \) for the longitude of the node precession (both the nodes of Jupiter and Saturn precess at the same rate, if measured relative to the *invariable plane*, defined as the plane orthogonal to their total angular momentum vector).

The dissipation of gas from the protoplanetary disk changes the gravitational potentials that the asteroids and planets feel, and hence changes the precession rates of their orbits. Given that the planets and asteroids are at different locations, they will be affected somewhat differently by this change of gravitational potential and consequently their precession rates will not change proportionally. It is therefore possible that secular resonances sweep through the asteroid belt as the gas dissipates. This means that every asteroid, whatever its location in the belt, first has orbital precession rates slower than the \( g_s \) and \( g_a \) frequencies of Jupiter and Saturn when there is a lot of gas in the disk, then enters resonance (\( g = g_s \) or \( g = g_a \)) when some appropriate fraction of the gas has been removed, and eventually is no longer in resonance (its orbital precession frequency being faster than those of the giant planets, i.e., \( g > g_a \)) after all the gas has disappeared. The same occurs for the asteroid’s nodal frequency \( s_p \) relative to the planetary frequency \( s_p \). This sweeping of perihelion and nodal secular resonances has the potential to excite the orbital eccentricities and inclinations of all asteroids.

This mechanism of asteroid excitation due to disk dissipation has been revisited with numerical simulations in Lemaitre and Dubru (1991), Lecar and Franklin (1997), Nagasawa et al. (2000, 2001, 2002), Petit et al. (2002), and finally by O’Brien et al. (2007). Nagasawa et al. (2000) found that of all the scenarios for gas depletion they studied (uniform depletion, inside-out, and outside-in), inside-out depletion of the nebula was most effective at exciting eccentricities and inclinations of asteroids throughout the main belt. However, they (unrealistically) assumed that the nebula coincided with the ecliptic plane. Protoplanetary disks can be warped, but they are typically aligned with the orbit of the locally dominant planet (Mouillet et al., 1997). Thus, there is no reason that the gaseous disk in the asteroid belt region was aligned with the current orbital plane of Earth (which was not yet formed). Almost certainly it was aligned with the orbits of the giant planets. Taking the invariable plane (the plane orthogonal to the total angular momentum of the solar system) as a proxy of the original orbital plane of Jupiter and Saturn, Nagasawa et al. (2001, 2002) found that the excitation of inclinations would be greatly diminished. Furthermore, since nebular gas in the inside-out depletion scenario would be removed from the asteroid belt region before the resonances swept through it, there would be no gas drag effect to help deplete material from the main-belt region.

The work of O’Brien et al. (2007) accounted for the fact that the giant planets should have had orbits significantly less inclined and eccentric than their current values when they were still embedded in the disk of gas, because of the strong damping that gas exerts on planets (Cresswell et al., 2008; Kley and Nelson, 2012). They concluded that secular resonance sweeping is effective at exciting eccentricities and inclinations to their current values only if gas is removed from the inside-out and very slowly, on a timescale of \(~20\) m.y. This gas-removal mode is very different from our current understanding of the photoevaporation process (Alexander et al., 2014), and inconsistent with observations suggesting that disks around solar-type stars have lifetimes of only \(1\)–\(10\) m.y., with an average of \(~3\) m.y. (e.g., Strom et al., 1993; Zuckerman et al., 1995; Kenyon and Hartmann, 1995; Haisch et al., 2001).

Earlier studies found that the final eccentricities of the asteroids are quite randomized because two perihelion secular resonances sweep the entire asteroid belt in sequence: first the resonance \( g = g_s \), then the resonance \( g = g_a \). The first resonance excites the eccentricities of the asteroids from zero to approximately the same value, but the second resonance, sweeping an already excited belt, can increase or decrease the eccentricity depending on the position of the perihelion of each asteroid at the time of the encounter with the resonance (Ward et al., 1976; Minton and Malhotra, 2011). O’Brien et al. (2007) found that when Jupiter and Saturn were on orbits initially closer together, as predicted by the Nice model (e.g., Tsiganis et al., 2005), the resonance with frequency \( g_a \) would only sweep part of the outer belt, leading to less randomization of eccentricities in the inner belt. All studies in which the mid-plane of the protoplanetary disk of gas coincides with the invariable plane of the solar system find that the final inclinations tend to have comparable values. This is because there is only one dominant frequency \( (s_p) \) in the precession of the nodes of Jupiter and Saturn and hence there is only one nodal secular resonance and no randomization of the final inclinations of the asteroids. Clearly, this is in contrast with the observations. For all these problems, the model of secular resonance sweeping during gas removal is no longer considered to be able to alone explain the excitation and depletion of the primordial asteroid belt.

An alternative model for the dynamical excitation of the asteroid belt was proposed by Ip (1987). In this model, putative planetary embryos are scattered out of the Jupiter region and cross the asteroid belt for some timescale before being ultimately dynamically ejected from the solar system. If the embryos are massive enough, their repeated crossing of the asteroid belt can excite and randomize the eccentricities and inclinations of the asteroids, through close encounters and secular effects. That scenario has been revisited by Petit et al. (1999), who found that, whatever the mass of the putative embryos, the resulting excitation in the asteroid belt ought to be very unbalanced. Excitation would be much stronger in the outer belt than in the inner belt (because the embryos come from Jupiter’s region) and it would be much stronger in eccentricity than in inclination. By contrast, the main asteroid belt shows no such trend (see Fig. 1). So, again, this model has since been abandoned. If massive embryos have been
scattered from Jupiter’s zone, they must have crossed the asteroid belt very briefly so that their limited effects could be completely overprinted by other processes, such as those discussed below.

3.2. Wetherill’s Model

The first comprehensive model of asteroid belt sculpting, which linked the evolution of the asteroid belt with the process of terrestrial planet formation, was that proposed by Wetherill (1992) and later simulated in a number of subsequent papers (e.g., Chambers and Wetherill, 1998; Petit et al., 2001, 2002; O’Brien et al., 2006, 2007). In this model, at the time gas was removed from the system, the protoplanetary disk interior to Jupiter consisted of a bimodal population of planetesimals and planetary embryos, the latter with masses comparable to those of the Moon or Mars. Numerical simulations show that, under the effect of the mutual perturbations among the embryos and the resonant perturbations from Jupiter, embryos are generally cleared from the asteroid belt region, whereas embryos collide with each other and build terrestrial planets inside 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 planet 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. A series of simulation snapshots demonstrating this process is shown in Fig. 3.

Whereas earlier simulations assumed that Jupiter and Saturn were originally on their current orbits, O’Brien et al. (2006, 2007) performed simulations with Jupiter and Saturn on the low-inclination, nearly circular orbits predicted in the Nice model. The resulting asteroids from a set of simulations with these initial conditions are shown in Fig. 4. Overall, the range of values compare well with those observed for the real asteroids, although the final inclination distribution is skewed toward large inclinations. The reason for this is that it is easier to excite a low-inclination asteroid to large eccentricity and remove it from the belt than it is for a high-inclination asteroid, because the encounter velocities with the embryos are slower and more effective in deflecting the low-inclination asteroid’s orbit. Also, with the giant planets on nearly circular orbits, it takes longer to clear embryos from the asteroid belt, allowing more time to excite asteroids to large inclinations.

As noted earlier, the surviving asteroids have their orbital semimajor axes displaced from their original values, as a result of the embryos’ gravitational scattering. O’Brien et al. (2007) found that the typical change in semimajor axis is on the order of 0.5 AU (comparable to earlier simulations), with a tail extending to 1–2 AU, and the semimajor axis can be either decreased or increased. This process can explain the partial mixing of taxonomic types. As shown in Fig. 2, the distribution of the S-type and C-type asteroids has a Gaussian-like shape, with a characteristic width of ~0.5 AU. Thus, if one postulates that all S-type asteroids originated from the vicinity of 2 AU and all C-type asteroids originated in the vicinity of 3 AU, Wetherill’s model explains the current distribution.

3.3. The Grand Tack Model

A more recent, alternative model to Wetherill’s is the so-called Grand Tack scenario, proposed in Walsh et al. (2011). Initially the Grand Tack scenario had not been developed to explain the asteroid belt, but to answer two questions left open by Wetherill’s model: Why is Mars so small relative to Earth? Why is Jupiter so far from the Sun despite planets having a tendency to migrate inward in protoplanetary disks? Nevertheless, this scenario has profound implications for the asteroid belt, as we discuss below.

The Grand Tack scenario is built on results from hydrodynamics simulations finding that Jupiter migrates toward 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 scenario postulates that Jupiter formed first. As long as the planet was basically alone, Saturn being too small to influence its dynamics, Jupiter migrated inward from its initial position (poorly constrained but estimated at ~3.5 AU) down to 1.5 AU. Then, when Saturn reached a mass close to its current one and an orbit close to that of Jupiter, Jupiter reversed migration direction (i.e., it “tacked,” hence the name of the model) and the pair of planets started to move outward. This outward migration continued until the final removal of gas in the disk, which the model assumes happened when Jupiter reached a distance of ~5.5 AU. The migration of the cores of giant planets is still not fully understood (see Kley and Nelson, 2012, for a review). Thus, the Grand Tack model comes in two “flavors.” In one, Saturn, while growing, migrates inward with Jupiter. In another, Saturn is stranded at a no-migration orbital radius until its mass exceeds 50 M⊕ (Bitsch et al., 2014), then it starts migrating inward and it catches Jupiter in resonance because it migrates faster. Both versions exist with and without Uranus and Neptune. All these variants are described in Walsh et al. (2011); the results are very similar in all these cases, which shows the robustness of the model, at least within the range of tested possibilities. The scheme presented in Fig. 5 has been developed in the framework of the first “flavor.”

Assuming that Jupiter formed at the snowline (a usual assumption to justify the large mass of its core and its fast formation), the planetesimals that formed inside its initial orbit should have been mostly dry. It is therefore reasonable to associate these planetesimals (whose distribution is sketched as a dashed area in Fig. 5) with the S-type asteroids and other even dryer bodies (enstatite-type, Earth precursors, etc.). During its inward migration, Jupiter penetrates into the disk of these planetesimals. In doing so, most planetesimals (and planetary embryos) are captured in mean-motion resonances...
with Jupiter and are pushed inward, increasing the mass density of the inner part of the disk. However, some 10% of the planetesimals are kicked outward by an encounter with Jupiter, reaching orbits located beyond Saturn that collectively have an orbital (a,e) distribution 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. 5) that are initially on circular orbits beyond the orbit of Saturn. These bodies, being formed beyond the snowline, should be rich in water ice and other volatile elements, and therefore it is again reasonable to associate them with C-type asteroids.

After reaching ~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)], Jupiter reverses its migration direction and begins its outward migration phase, during which the giant planets encounter the scattered S-type disk, and then also the primitive C-type disk. Some of the bodies in both populations are scattered inward, reach the asteroid-belt region, and are implanted there as Jupiter moves out of it.
The final orbits of the planetesimals, at the end of the outward migration phase, are shown in Fig. 6. A larger dot size is used to highlight the planetesimals trapped in the asteroid belt region and distinguish them from those in the inner solar system or at eccentricities too large to be in the asteroid belt. Notice that the asteroid belt is not empty, although it has been strongly depleted (by a factor of several hundred relative to its initial population). This result is not trivial. One could have expected that Jupiter migrating through the asteroid belt twice (first inward then outward) would have completely removed the asteroid population, invalidating the Grand Tack scenario. The eccentricities and the inclinations of the particles in the asteroid belt are excited and randomized. The S-type particles (black) are found predominantly in the inner part of the belt and the C-type particles (gray) in the outer part, but there is a wide overlapping region where both are present. This is qualitatively consistent with what is observed.

As discussed above, the Grand Tack scenario solves open problems in Wetherill's model. The small mass of Mars is explained as a result of the disk of the remaining solid material being truncated at ~1 AU (Hansen, 2009; Walsh et al., 2011). Fischer and Ciesla (2014) reported that they could obtain a small-mass Mars in a few percent of simulations conducted in the framework of Wetherill’s model. However, the rest of the planetary system in these simulations does not resemble the real terrestrial planet system (Jacobson and Walsh, 2015). For instance, another massive planet is formed in the Mars region or beyond. The outward migration of Jupiter explains why the giant planets in our solar system are so far from the Sun, whereas most giant planets found so far around other stars are located at 1–2 AU. For all these reasons, one can consider the Grand Tack model more as an improvement of Wetherill’s model than an alternative, because it is built in the same spirit of linking the asteroid belt sculpting to the evolution of the rest of the solar system (terrestrial planet formation and giant-planet migration, the latter of which was still unknown during Wetherill’s time).

It is nevertheless interesting to compare the Grand Tack model and Wetherill’s model on the basis of the final asteroid belts that they produce. Comparing Fig. 6 with Fig. 4, it is apparent that the Grand Tack model provides a better inclination distribution, more uniform than Wetherill’s, but it produces a worse eccentricity distribution, which is now more skewed toward the upper eccentricity boundary of the asteroid belt.

As we will see in section 3.5, however, the eccentricity distribution can be remodeled somewhat during a later evolutionary phase of the solar system. This is also partially true for the inclination distribution. So, for what concerns the eccentricity and inclination distributions, one might declare a tie in the competition between the two models.

![Fig. 5. 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 planets migrate inward, then, when Saturn reaches its current mass, they move outward. The dashed and dotted areas schematize the (a,e) distributions of S-type and C-type asteroids respectively. The dashed and dotted arrows in the lower panel illustrate the injection of scattered S-type and C-type asteroids into the asteroid belt during the final phase of outward migration of the planets.](image)

![Fig. 6. Final semimajor axis, eccentricity, and inclination distribution of bodies surviving the inward and outward migration of Jupiter and Saturn. The black particles were originally placed inside the initial orbit of Jupiter and the gray particles outside the initial orbit of Saturn. The particles finally trapped in the asteroid belt are depicted with larger symbols than the others. The dashed curve in the lower panel shows the approximate boundaries of the asteroid belt inward of the 2:1 resonance with Jupiter. This final distribution was achieved in the simulations of Walsh et al. (2011) accounting only for Jupiter and Saturn (i.e., not including Uranus and Neptune) moving together in the 2:3 resonance, as shown in Fig. 5.](image)
The Grand Tack model makes it conceptually easier to understand the significant differences between S-type and C-type asteroids and their respective presumed daughter populations: the ordinary and carbonaceous chondrites. In fact, in the Grand Tack model these two populations are sourced from clearly distinct reservoirs on either side of the snowline. Instead, in Wetherill’s model these bodies would have formed just at the two ends of the asteroid belt, so less than 1 AU apart. Despite such a vast difference in predicted formation locations for these two populations, the debate is open. Some authors (e.g., Alexander et al., 2012) think that bodies formed in the giant-planet region would be much more similar to comets than to asteroids, while others (Gounelle et al., 2008) argue that there is a continuum between C-type asteroids and comets and a clear cleavage of physical properties between ordinary and carbonaceous chondrites. We review the available cosmochemical constraints and their uncertain compatibility with the model in section 3.4.

A clear distinction between the Grand Tack model and Wetherill’s model is that the former provides a faster and more drastic depletion of the asteroid belt. This point is illustrated in Fig. 7, which shows the fraction of the initial asteroid population that is in the main-belt region at any time.

The Grand Tack scenario depletes the asteroid belt down to 0.3%, and does so basically in 0.1 m.y. Assuming that the final asteroid belt consisted of one current asteroid belt mass in S-type asteroids and three current asteroid belt masses in C-type asteroids (the reason for 4× more total mass in the asteroid belt will be clarified in section 3.6), this implies that the asteroid belt at t = 0 should have contained 0.6 M⊕ in planetesimals (the rest in embryos). Also, a calculation of the collision probability of the asteroids as a function of time (both among each other and with the planetesimals outside the asteroid belt) shows that the integrated collisional activity suffered by the surviving asteroids during the first 200 m.y. would not exceed the equivalent of 4 G.y. in the current population. Thus, assuming that the exceeding factor of 4 in the asteroid population is lost within the next 500 m.y. (see sections 3.5 and 3.6), the integrated collisional activity of asteroids throughout the entire solar system age would probably remain within the 10-G.y. constraint described in section 2. In contrast, Wetherill’s model depletes the asteroid belt on a timescale of 100 m.y. Also, about 2–3% of the initial population remains in the belt at the end. Thus, to be consistent with constraints on the current population and its integrated collisional activity, the initial mass in planetesimals in the asteroid-belt region should have been no larger than 200× the current asteroid-belt mass, or less than one Mars mass (Botke et al., 2005b).

3.4. Are Cosmochemical Constraints Consistent with the Grand Tack Model?

The Grand Tack model predicts that C-type asteroids have been implanted into the asteroid belt from the giant-planet region. Is this supported or refuted by cosmochemical evidence?

Although there is a spread in values, the D/H ratios of carbonaceous chondrites (with the exception of CR chondrites) are a good match to Earth’s water (Alexander et al., 2012). Oort cloud comets are usually considered to have formed in the giant-planet region (e.g., Dones et al., 2004). The D/H ratio was measured for the water from seven Oort cloud comets (see Bockelée-Morvan et al., 2012, and references within). All but one [Comet 153P/Ikeya-Zhang (Biver et al., 2006)] have water D/H ratios about twice as high as chondritic. This prompted Yang et al. (2013) to develop a model where the D/H ratio of ice in the giant-planet region is high. However, Brasser et al. (2007) showed that comet-sized bodies could not be scattered from the giant-planet region into the Oort cloud in the presence of gas drag (i.e., when the giant planets formed), and Brasser and Morbidelli (2013) demonstrated that the Oort cloud population is consistent with an origin from the primordial transneptunian disk at a later time. The recent measurement (Altwegg et al., 2014) of a high D/H ratio for the ice of Comet 67P/Churyumov-Gerasimenko, which comes from the Kuiper belt, supports this conclusion by showing that there is no systematic difference between Oort cloud comets and Kuiper belt comets. Care should therefore be taken in using Oort cloud comets as indicators of the D/H ratio in the giant-planet region. Conflicting indications on the local D/H ratio come from the analysis of Saturn’s moons. Enceladus’ D/H ratio is roughly twice Earth’s (Waite et al., 2009), but Titan’s D/H ratio is Earth-like (Coustenis et al., 2008; Abbas et al., 2010; Nixon et al., 2012).

Alexander et al. (2012) also noticed a correlation between D/H and C/H in meteorites and interpreted it as evidence
for an isotopic exchange between pristine ice and organic matter within the parent bodies of carbonaceous chondrites. From this consideration, they argued that the original water reservoir of carbonaceous asteroids had a D/H ratio lower than Titan, Enceladus, or any comet, again making asteroids distinct from bodies formed in the giant-planet region and beyond. However, a reservoir of pristine ice has never been observed; the fact that Earth’s water and other volatiles are in chondritic proportion (Marty, 2012; however, see Halliday, 2013) means that carbonaceous chondrites — wherever they formed — reached their current D/H ratios very quickly, before delivering volatiles to Earth. It is also possible that the D/H ratio measured for comets and satellites might have been the result of a similar rapid exchange between a pristine ice and the organic matter.

Another isotopic constraint comes from the nitrogen isotope ratio. Comets seem to have a rather uniform $^{15}\text{N}/^{14}\text{N}$ ratio (Rousselot et al., 2014). Even the comets with a chondritic D/H ratio [e.g., Comet Hartley 2 (Hartog et al., 2011)] have a nonchondritic $^{15}\text{N}/^{14}\text{N}$ ratio (Meech et al., 2011). The $^{15}\text{N}/^{14}\text{N}$ ratio, however, is only measured in HCN or NH$_3$, never in molecular nitrogen (N$_2$). Titan has a cometary $^{15}\text{N}/^{14}\text{N}$ ratio as well [in this case measured in N$_2$ (Niemann et al., 2010)]. Here, again, a few caveats are in order. First, it is difficult to relate the composition of a satellite, born from a circumplanetary disk with its own thermal and chemical evolution, to the composition of bodies born at the same solar distance but on heliocentric orbits. Second, it is unclear whether any comets for which isotope ratios have been measured originate from the giant-planet region, as opposed to the transneptunian disk (Brasser and Morbidelli, 2013). We also point out that hot spots with large $^{15}\text{N}/^{14}\text{N}$ ratios are found in primitive meteorites (Busemann et al., 2006).

Arguments in favor of an isotopic similarity between carbonaceous chondrites and comets come from the analysis of micrometeorites. Most micrometeorites (particles of ~100 µm collected in the Antarctic ice) have chondritic isotope ratios for H, C, N, and O [with the exception of ultracarbonaceous micrometeorites, which have a large D/H ratio, comparable to that in the organics of some chondrites, but which constitute only a minority of the micrometeorite population (Duprat et al., 2010)]. Yet according to the best available model of the dust distribution in the inner solar system (Nesvorny et al., 2010), which is compelling given that it fits the zodiacal light observations almost perfectly, most of the dust accreted by Earth should be cometary, even when the entry velocity bias is taken into account. Similarly, from orbital considerations, Gounelle et al. (2006) concluded that the CI meteorite Orgueil is a piece of a comet. Compelling evidence for a continuum between chondrites and comets also comes from the examination of Comet Wild 2 particles returned to Earth by the Stardust mission (e.g., Zolensky et al., 2006, 2008; Ishii et al., 2008; Nakamura et al., 2008).

These considerations suggest that if one looks at their rocky components, comets and carbonaceous asteroids are very similar from a compositional and isotopic point of view, if not in fact indistinguishable.

Finally, it has been argued that if the parent bodies of carbonaceous chondrites had accreted among the giant planets, they would have contained ~50% water by mass. Instead, the limited hydrous alteration in carbonaceous meteorites suggests that only about 10% of the original mass was in water (A. N. Krot, personal communication, 2014; but see Alexander et al., 2010). However, a body’s original water content cannot easily be estimated from its aqueous alteration. Even if alteration is complete, there is a finite amount of water that the clays can hold in their structures. Thus, the carbonaceous parent bodies may have been more water-rich than their alteration seems to imply. In fact, the discoveries of main-belt comets releasing dust at each perihelion passage (Hsieh and Jewitt, 2006), of water ice on asteroids Themis (Campins et al., 2010; Rivkin and Emery, 2010) and Cybele (Licandro et al., 2011; Hargrove et al., 2012), and of vapor plumes on Ceres (Kuppers et al., 2014) show that C-type asteroids are more rich in water than their meteorite counterpart seems to suggest, supporting the idea that they might have formed near or beyond the snowline. Meteorites may simply represent rocky fragments of bodies that were far wetter/icier.

Clearly, the debate on whether carbonaceous asteroids really come from the giant-planet region as predicted by the Grand Tack model is wide open. More data are needed from a broader population of comets. The investigation of main-belt comets, both remote and in situ, and the Dawn mission at Ceres will be key to elucidating the real ice content of carbonaceous asteroids and their relationship with classic “comets.”

3.5. The Nice Model: A Second Phase of Excitation and Depletion for the Asteroid Belt

Figure 7 seems to suggest that after ~100 m.y. the asteroid belt had basically reached its final state. However, at this time the orbits of the giant planets were probably still not the current ones. In fact, the giant planets are expected to have emerged from the gas-disk phase in a compact and resonant configuration as a result of their migration in the gas-dominated disk. This is true not only in the Grand Tack model, but in any model where Jupiter is refrained from migrating inside ~5 AU by whatever mechanism (Morbidelli et al., 2007).

The transition of the giant planets from this early configuration to the current configuration is expected to have happened via an orbital instability, driven by the interaction of the planets with a massive disk of planetesimals located from a few astronomical units beyond the original orbit of Neptune to about 30 AU (Morbidelli et al., 2007; Batygin and Brown, 2010; Levison et al., 2011; Nesvorny, 2011; Batygin et al., 2012; Nesvorny and Morbidelli, 2012; see also a precursor work by Thommes et al., 1999). In essence, the planetesimals disturbed the orbits of the giant planets and, as soon as two planets fell off resonance, the entire system became unstable. In the simulations, the instability can occur early (e.g., Morbidelli et al., 2007) or late (Levison et al., 2011), depending on the location of the inner edge of the transneptunian disk.
Constraints suggest that in the real solar system the instability occurred relatively late, probably around 4.1 G.y. ago (namely 450 m.y. after gas removal). These constraints come primarily from the Moon. Dating lunar impact basins is difficult, because it is not clear which samples are related to which basin (e.g., Norman and Nemchin, 2014). Nevertheless, it is clear that several impact basins, probably a dozen, formed in the 4.1–3.8-Ga period (see Fassett and Minton, 2013, for a review). Numerical tests demonstrate that these late basins (even just Imbrium and Orientale, whose young ages are undisputed) are unlikely to have been produced by a declining population of planetesimals, left over from the terrestrial planet accretion process, because of their short dynamical and collisional lifetimes (Botke et al., 2007).

There is also a surge in lunar rock impact ages ~4 G.y. ago, which contrasts with a paucity of impact ages between 4.4 and 4.2 Ga (Cohen et al., 2005). This is difficult to explain if the bombardment had been caused by a population of leftover planetesimals slowly declining over time. The situation is very similar for the bombardment of asteroids, with meteorites showing a surge in impact ages at 4.1 Ga and a paucity of ages between 4.2 and 4.4 Ga (Marchi et al., 2013). Meteorites also show many impact ages near 4.5 Ga, demonstrating that the apparent lack of events in the 4.2–4.4-Ga interval is not due to clock resetting processes. All these constraints strongly suggest the appearance of a new generation of projectiles in the inner solar system about 4.1 G.y. ago, which argues that either a very big asteroid broke up at that time (Cuk, 2012) (but such a breakup is very unlikely from collision probability arguments and we do not see any remnant asteroid family supporting this hypothesis), or that the dynamical instability of the giant planets occurred at that time, partially destabilizing small-body reservoirs that had remained stable until then.

Other constraints pointing to the late instability of the giant planets come from the outer solar system. If the planets had become unstable at the disappearance of the gas in the disk, presumably the Sun would still have been in a stellar cluster and consequently the Oort cloud would have formed more tightly bound to the Sun than it is thought to be from the orbital distribution of long-period comets (Brasser et al., 2008, 2012). Also, the impact basins on lapetus (a satellite of Saturn) have topographies that have relaxed by 25% or less, which argues that they formed in a very viscous lithosphere; according to models of the thermal evolution of the satellite, these basins can not have formed earlier than 200 m.y. after the beginning of the solar system (Robuchon et al., 2011).

For all these reasons, it is appropriate to discuss the consequences of the giant-planet instability on the asteroid belt, after the events described by the Grand Tack or Wetherill’s models. In fact, it is important to realize that the model of giant-planet instability (often called the Nice model) is not an alternative to the models described before on the early evolution of the asteroid belt; instead it is a model of the subsequent evolution.

The phase of giant-planet instability is very chaotic and therefore a variety of orbital evolutions are possible. Nevertheless, the planetary evolutions can be grouped in two categories. In the first category, Saturn, Uranus, and Neptune have close encounters with each other, but none of them have encounters with Jupiter. Saturn typically scatters either Uranus or Neptune outward and thus it recoils inward. As a result, Uranus and Neptune acquire large eccentricity orbits that cross the transneptunian disk. The dispersal of the planetesimal disk damp the eccentricities of the planets by dynamical friction and drives the planets’ divergent migration in semimajor axis (Tsiganis et al., 2005). Thus, the planets achieve stable orbits that are well separated from each other. The orbital separation between Jupiter and Saturn first decreases, when Saturn recoils, and then increases due to planetesimal-driven migration. The timescale for the latter is typically ~10 m.y. The slow separation between the two major planets of the solar system drives a slow migration of secular resonances across the asteroid belt (Minton and Malhotra, 2009, 2011) and the terrestrial planet region (Brasser et al., 2009).

The jumping-Jupiter evolution also explains the capture of the irregular satellites of Jupiter with an orbital distribution similar to those of the irregular satellites of the other giant planets (Nesvorný et al., 2007, 2014). It can also explain the capture of Jupiter’s Trojans in uneven proportions around the L4 and L5 Lagrangian points (Nesvorný et al., 2013; see also the chapter by Emery et al. in this volume). So far, no other model is capable of achieving these results. For all these reasons, simulated solar system evolutions these days are required to show a jumping-Jupiter evolution to be declared successful (Nesvorný and Morbidelli, 2012).

Although it avoids secular resonances sweeping across the asteroid belt, a jumping-Jupiter evolution is not without consequences for the asteroids. The sudden change in the eccentricity of Jupiter (from an initial basically circular orbit, like that observed in hydrodynamical simulations of the four giant planets evolving in the gaseous protoplanetary disk, to one with current eccentricity) changes the forced eccentricity felt by the asteroids in their secular evolution. Consequently, the proper eccentricities of the asteroids are changed. Depending on the value of the longitude of
perihelion when the forced eccentricity changes, the proper eccentricity of an asteroid can decrease or increase. Roughly, 50% of the asteroids are kicked to larger eccentricities and therefore are removed from the asteroid belt. The remaining 50% of the asteroids have their eccentricities reduced. This can potentially reconcile the eccentricity distribution of asteroids at the end of the Grand Tack evolution with the current distribution. Indeed, Minton and Malhotra (2011) showed that the current eccentricity distribution can be achieved starting from a primordial distribution peaked at $e \sim 0.3$, similar to that produced by the Grand Tack evolution of Jupiter (Fig. 6). They obtained this result using secular resonance sweeping, but the basic result should hold also for a sudden enhancement of Jupiter’s eccentricity. Nevertheless, specific numerical simulations have never been done to demonstrate that the eccentricity distribution of asteroids at the end of the Grand Tack model can be transformed into the current distribution via the jumping-Jupiter evolution [at the time of this writing, this work is ongoing and seems to give positive results (Deienno and Gomes, personal communication)]. The jump in Jupiter’s inclination due to its encounter(s) with another planet should also have partially reshuffled the asteroid inclination distribution, possibly reconciling the final distribution in Wetherill’s model (see Fig. 4) with the current one. However, the effects on the inclination during Jupiter’s jump seem less pronounced than those on the eccentricity (Morbidelli et al., 2010).

The current inner edge of the asteroid belt is marked by the presence of the secular resonance between the precession frequency of the perihelion of an asteroid and the $g_2$ planetary frequency. The resonance makes unstable all objects inside 2.1 AU at low to moderate inclinations, which truncates the belt at its current edge. But before the impulsive separation between the inner orbits of Jupiter and Saturn, this resonance was much weaker (because the $g_2$ mode was less excited in the planetary system, the giant-planet orbits being more circular) and located away from the asteroid belt. Thus, in principle the asteroid belt might have extended closer to Mars before the giant-planet-instability event. Bottke et al. (2012) showed that the destabilization of this extended belt — or “E belt” — could have dominated the formation of impact basins on the Moon, producing 12–15 basins over a 400-m.y. interval. Given the age of the Orientale basin (usually estimated at 3.7–3.8 Ga), this implies that the giant-planet instability and the destabilization of the E belt happened 4.1–4.2 G.y. ago and was responsible for the production of the last dozen lunar basins, known as the Nectarian and post-Nectarian basins. Earlier basins and craters would have to have come from other sources, such as the planetesimals leftover from the terrestrial planet-formation process. The existence of two populations of projectiles, namely the leftover planetesimals dominating the bombardment at early times and the E-belt asteroids dominating the impact rate at a later epoch, should have produced a sawtooth-shaped bombardment history of the Moon (Morbidelli et al., 2012). The E belt should also have caused a long, slowly decaying tail of Chicxulub-sized impacts on Earth, possibly continuing until ~2 G.y. ago. Evidence for this long tail in the time-distribution of impacts is provided by the existence of terrestrial impact spherule beds, which are globally distributed ejecta layers created by the formation of Chicxulub-sized or larger craters (Johnson and Melosh, 2012): 10, 4, and 1 of these beds have been found on well-preserved, non-metamorphosed terrains between 3.23 and 3.47 Ga, 2.49 and 2.63 Ga, and 1.7 and 2.1 Ga, respectively (Simonson and Glass, 2004; Bottke et al., 2012).

Moreover, the escape to high-eccentricity orbits of bodies from the main-belt and E-belt regions produced a spike in the impact velocities on main-belt asteroids at the time of the giant-planet instability. Thus, although the impact frequency on asteroids decreased with the depletion of 50% of the main-belt population and 100% of the E-belt population, the production of impact melt on asteroids increased during this event because melt production is very sensitive to impact velocities (Marchi et al., 2013). For this reason, the impact ages of meteorites show a spike at 4.1 Ga like

![Figure 8](image.png)

**Fig. 8.** A comparison between the final (a,i) distribution of asteroids if (a) Jupiter and Saturn migrate slowly away from each other or (b) jump due to them scattering an ice giant planet. In the first case the orbital distribution is inconsistent with that observed, while in the second case it is. From Morbidelli et al. (2010).
the lunar rocks, although for the latter this is due to a surge in the impact rate.

A final consequence of the giant-planet instability on the asteroid belt is the capture into its outer region of planetesimals from the transneptunian disk (Levison et al., 2009). Because Jupiter Trojans are captured in the same event, these last captured asteroids should have had the same source as the Trojans and therefore they should be mostly taxonomic D- and P-types. The probability of capture in the asteroid belt is nevertheless small, so it is unlikely that an object as big as Ceres was trapped from the transneptunian region in this event.

3.6. After the Giant-Planet Instability

After the giant-planet instability the orbits of all planets, giants and terrestrial, are similar to the current ones (within the range of semimajor axis, eccentricity, and inclination oscillation provided by the current dynamical evolution). Thus, the asteroid belt has finished evolving substantially under the effect of external events such as giant-planet migration or instability.

The asteroid belt thus entered into the current era of evolution. Asteroids became depleted at the locations of unstable resonances (mean-motion and secular) on timescales that varied from resonance to resonance. In this way, the asteroid belt acquired its current final orbital structure. In this process, it is likely that another ~50% of the asteroids were removed from the belt, most of them during the first 100 m.y. after the giant-planet instability (Minton and Malhotra, 2010). Combining this 50% with the 50% loss during the instability itself is the reason that we require that the primordial depletion event (Wetherill’s model or Grand Tack) left a population of asteroids in the belt that was about 4× the current one.

With the depletion of unstable resonances, the asteroid belt would have become an extraordinarily boring place from a dynamical point of view. Fortunately, collisional breakup events keep refreshing the asteroid population, generating dynamical families very rich in small objects, while non-gravitational forces, mostly the Yarkovsky effect (Bottke et al., 2006), cause small asteroids to drift in semimajor axis, eventually supplying new bodies to the unstable resonances. This combination of collisional activity and nongravitational forces allow the main asteroid belt to resupply and sustain in a quasi-steady state the intrinsically unstable population of near Earth objects. But this is the subject of another chapter.

4. CONCLUSIONS AND IMPLICATIONS

In this chapter, we have reviewed our current understanding of the evolution of the asteroid belt, from a massive and dynamically quiet disk of planetesimals to its current state, which is so complex and rich from the points of view of both its dynamical and physical structures.

According to this understanding, the asteroid population mainly evolved in two steps. There was an early event of strong dynamical excitation and asteroid removal, which left about 4× the current asteroid population on orbits with a wide range of eccentricities and inclinations. This event may have been due to the self-stirring of a population of planetary embryos resident in the asteroid belt (Wetherill, 1992), or to the migration of Jupiter through the asteroid belt [the Grand Tack scenario (Walsh et al., 2011)]. The second step occurred later, possibly as late as 4.1 G.y. ago or ~400 m.y. after the removal of gas from the protoplanetary disk. At that time, the asteroid belt underwent a second dynamical excitation and depletion when the giant planets became temporarily unstable and their orbits evolved from an initial resonant and compact configuration to the current configuration. During this second event, the asteroid belt lost about 50% of its members. After this second event, the asteroid belt structure settled down with the progressive depletion at unstable resonances with the giant and terrestrial planets. Another 50% of the asteroid population was lost in this process, mostly during the subsequent 100 m.y.

If the first evolutionary step was due to the Grand Tack migration of Jupiter, we expect that S-type asteroids formed more or less in situ (2–3-AU region); the C-type asteroids formed in the giant-planet region (roughly 3–15 AU), and the P- and D-type asteroids formed beyond the initial location of Neptune (roughly 15–30 AU). The hot population of the Kuiper belt, the scattered disk, and the Oort cloud would also derive from the same transneptunian disk (Levison et al., 2008; Brasser and Morbidelli, 2013). There is a growing consensus that the cold Kuiper belt (42–45 AU) is primordial and born in situ (Pettit et al., 2011; Parker et al., 2011; Batygin et al., 2011; Fraser et al., 2014). Thus, the cold Kuiper belt objects should not have any correspondent in the asteroid belt.

If instead the first step was due to the self-stirring of resident embryos as in Wetherill’s model, we expect that S-type asteroids formed in the inner part of the belt, C-type asteroids in the outer part, and no asteroids sample the planetesimal population in the giant-planet region. The origin of P- and D-type asteroids would be the same as above. Thus, deciding which of these two models is preferable requires a better understanding of the nature of C-type asteroids and their water content, the similarities and differences between them and comets, and among comets themselves. This may not be an easy task. The population of main-belt comets (asteroids showing cometary activity such as 133P/Elst-Pizarro or 238P/Read) and their relationship with the parent bodies of carbonaceous chondrites is key in this respect. If it turns out that main-belt comets are consistent with carbonaceous chondrites in terms of isotope composition (mostly for H, N, and O), then this will argue that carbonaceous chondrites are just the rocky counterpart of bodies much richer in water/ice than meteorites themselves. This would imply that C-type asteroids formed beyond the snowline, thus presumably in the vicinity of the giant planets. If instead the main-belt comets...
are not related to carbonaceous chondrites, but are more similar to comets from their isotope composition (it should be noted that even though Comet Hartley 2 has a chondritic water D/H ratio, it has a nonchondritic 14N/15N ratio), then this would argue for their injection in the belt from the cometary disk and would suggest that the parent bodies of carbonaceous chondrites are not so water-rich, and therefore formed somewhat closer to the Sun than the snowline.

Whatever the preferred scenario for the first depletion and excitation of the asteroid belt, it is clear that the asteroid population must have suffered in the first hundreds of millions of years as much collisional evolution as over the last 4 G.y. However, all asteroid families formed during the early times are not identifiable today because the dynamical excitation events dispersed them (and possibly depleted them) too severely. The presence of metallic asteroids not associated with a family of objects of basaltic or dunitic nature, as well as the existence of rogue basaltic asteroids such as 1459 Magnya, should therefore not come as a surprise. The only families that are preserved are those that formed after the last giant-planet-instability event and have not been made unrecognizable by subsequent collisional evolution and Yarkovsky drift; thus they are either relatively young or large.

In this chapter, we have also examined several other asteroid excitation and depletion scenarios, most of which have serious difficulties in reconciling their predicted outcomes with observations. We have done this not just for historical completeness, but also to illustrate the critical constraints on putative alternative scenarios of solar system evolution. For instance, numerous studies on the possible in situ formation of extrasolar super-Earths close to their host stars assume a large pileup of drifting material of various sizes, from grains to small-mass embryos, in the inner part of the protoplanetary disks (Hansen, 2014; Chatterjee and Tan, 2014; Boley et al., 2014). By analogy, these models could be used to suggest that the outer edge of the planetesimal disk at 1 AU, required to form a small Mars, was due to the same phenomenon rather than to the Grand Tack migration of Jupiter. However, from what we reported in this chapter, we think that the asteroid belt rules out this possibility. In fact, the inward migration of small planetesimals (due to gas drag) and large embryos (due to disk tides) could explain the pileup of solid mass inside 1 AU and the mass deficit of the asteroid belt, but not the asteroids’ orbital distribution (Izidoro et al., 2015). In the absence of the Grand Tack migration of Jupiter, we showed in section 3 that the only mechanism that could give the belt an orbital structure similar to that observed is Wetherill’s model of mutual scattering of resident embryos. But if this was the case, then the mass distribution could not be concentrated within 1 AU because a massive population of embryos is required in the main-belt region. Thus, at the current state of knowledge (which may change in the future), only the Grand Tack scenario seems able to explain the required mass concentration to make a small Mars.

In summary, the asteroid belt remains the population of choice to test old, current, and future models of solar system evolution.

**Acknowledgments.** A.M. was supported by the European Research Council (ERC) Advanced Grant ACCRETE (contract number 290568).

**REFERENCES**


