Asteroid collisional evolution studies are aimed at understanding how collisions have shaped observed features of the asteroid population in order to further our understanding of the formation and evolution of our solar system. We review progress in developing more realistic collisional scaling laws, the effects of relaxing over-simplifying assumptions used in earlier collisional evolution studies, and the implications of including observables, such as collisionally produced families, on constraining the collisional history of main-belt asteroids. Also, collisional studies are extended to include Jupiter Trojans and the Hilda population. Results from collisional evolution models strongly suggest that the mass of main-belt asteroids was only modestly larger, by up to a factor of 5 or so, at the time that the present collisionally erosive environment was established, presumably early in solar system history. Major problems remain in identifying the appropriate scaling algorithm for determining the threshold for catastrophic disruption as well as understanding the resulting size and velocity distribution of fragments. Dynamical effects need to be combined with collisional simulations in order to understand the structure of the small asteroid size distribution.

1. INTRODUCTION

The importance of collisions in shaping the present asteroid belt has been recognized for nearly 50 years, following the pioneering work of Piotrowski (1953) on the frequency of asteroidal collisions. Since that time, increasingly sophisticated tools have been developed for tracing the collisional history of asteroid populations and understanding the role of collisions in producing observed features of various small-body populations. Here we review the present state of understanding of asteroid collisional evolution, concentrating on progress since the Asteroids II volume.

Observations relevant to collisional processes are described in section 2, including recent work on extending the asteroid size distribution to subkilometer sizes. Determining the size distribution is a key constraint for collisional models and is a major discriminant among the various scaling laws proposed for describing collisional outcomes. In section 3 we summarize work on scaling laws that describes the outcome of collisions involving bodies of various sizes hitting at speeds of multikilometers per second. Understanding collisional outcomes for asteroid-sized bodies is fundamental to understanding asteroid collisional history; however, a consensus has not yet developed among workers as to the appropriate methodology to go from laboratory-scale experiments to giant collisions involving bodies to hundreds of kilometers in size. Gravity becomes significant in the asteroid size range between subkilometer-sized bodies and the ~850-km-diameter Ceres and thus adds to the complexity of understanding collisional outcomes. The effects of size-dependent collisional outcomes as well as boundary conditions on the overall population distribution are active topics of research. As discussed in section 3, the existence of a small size cutoff in a collisionally interacting population can produce a wavelike structure, a feature not recognized in earlier work on collisional evolution.

Collisional evolution of Trojans and Hildas has been studied in depth in recent years, as reviewed in section 3. These populations are particularly interesting to study because they are dynamically isolated from much of the main asteroid belt, although they do collisionally interact with each other. Additionally, these populations are dynamically stable, but collisions can eject fragments into unstable orbits that are then lost from the system. Comparing the col-
lisional evolution of these populations with each other provides insight to features common to the collisional evolution of small-body populations. Another collisionally evolved population discovered recently is the Kuiper Belt, a vast population of remnant planetesimals orbiting beyond Neptune. Kuiper Belt objects are thought to share many features with main-belt asteroids, including a significant role for collisions in shaping the present Kuiper Belt. However, this chapter will be restricted to the role of collisions on asteroids in the inner solar system, and the interested reader is referred to Farinella et al. (2000) for a comprehensive review of the role of collisions in the Kuiper Belt.

Finally, section 5 gives our perspective on the major outstanding problems in the field of asteroid collisional studies. We expect that by the time of the publication of Asteroids IV, most of these will be solved and an equally challenging set of questions will be posed for the next decade.

2. OBSERVATIONAL CONSTRAINTS

The litmus test for models of the collisional history of asteroids is how well they match the observed characteristics of the present asteroid belt, i.e., those features of the belt that are the product of 4.5 b.y. of collisional history. So, what are the collisionally produced observables in the asteroid belt and other small body populations? First is the size-frequency distribution; Fig. 1 shows the inferred size-frequency distribution for (a) main-belt asteroids, (b) Trojans, and (c) Hildas. Throughout this chapter we use the cumulative diameter distribution, i.e., the number of bodies larger than diameter D, to describe the size distribution of populations. This relationship is expressed mathematically as \( N(>D) = K D^{-b} \) where \( b \) is the cumulative diameter population index. We also use the incremental size distribution for clarity in some cases; however, the population index is numerically the same in these two formulations. An excellent discussion of the various power laws used to describe small-body populations is given by Colwell (1994). It is widely believed that the largest asteroids (although where “largest” begins is a debated question) are primordial objects whose sizes have not been significantly altered by collisions; only the smallest size end of the population is collisionally evolved.

The so-called “bump” in the main-belt asteroid size distribution centered at about 100 km diameter has been explained in three different ways: (1) Anders (1965) proposed that this was the signature of collisional modification of an initial Gaussian-like planetesimal size distribution. (2) Davis et al. (1979, 1984) found that this feature marked the transition from the gravity strength regime to the strength dominant size regime and hence was a product of collisional physics. (3) Durda et al. (1998) argued that this feature was a secondary “bump” produced by the wave from the strength-gravity transition that occurs at much smaller sizes (see section 3).

While there is structure in the small asteroid size distribution, a power-law fit to the cumulative absolute magnitude distribution of the Palomar Leiden Survey (PLS) (Kresák, 1977; Van Houten et al., 1970) gives a slope of 1.95 over the size range larger than a few kilometers (\( D \geq 2–5 \) km). Figure 2 shows various recent estimates of the small end of the main-belt size small distribution (see also Jedicke et al., 2002).

It is worth remarking that several lines of evidence point to the existence of a variable index in the size distribution of the small asteroids. For main-belt asteroids smaller than about 20 km, the PLS (Van Houten et al., 1970) found variable indexes. Also, Cellino et al. (1991) analyzed IRAS data for different zones of the asteroid belt and different size ranges (but always exceeding a few tens of kilometers, to avoid discovery biases) and found indexes ranging from \(-2\) to \(-4\). Similar variability occurs for objects with orbits crossing those of the inner planets, as suggested by the lunar and martian cratering record (Grün et al., 1985; Strom et al., 1992), the in situ dust detection experiments carried out on spacecraft (Grün et al., 1985; Love and Brownlee, 1993), and telescopic searches for small Earth-approaching asteroids (Ceplecha, 1992; Rabinowitz, 1993). Analyses of the cratering record on the surface of 951 Gaspra, imaged in 1991 by the Galileo probe, have indicated that a cumulative size distribution index of \(-2.7 \pm 0.5\) over the size range 0.4–1.5 km (Belton et al., 1992), while for Ida, the index varies from \(-3.3\) to \(-3.5\) for crater diameters \( \geq 1 \) km (Belton et al., 1994). The craters were produced by projectiles (~20 to 200 m across) that have cratered the surfaces of Gaspra and Ida.

A powerful constraint on the asteroid collisional history is the existence today of the basaltic crust of Vesta, which, if it is the ultimate source of the eucrite meteorites as is widely believed, dates to the earliest era of the solar system some 4.54 b.y. ago. Any collisional model must preserve this 25–40-km-thick crust (Gaffey et al., 1993) during the collisional bombardment over solar system history (Davis et al., 1984). Recent Hubble Space Telescope (HST) observations

![Fig. 1](image-url)
of Vesta revealed the existence of a ~450-km-diameter basin thought to be formed by the impact of a ~40-km-diameter projectile (Marzari et al., 1996; Asphaug, 1997). This impact, presumably the largest since Vesta’s crust formed, provides a very specific constraint on Vesta’s collisional history. Yet this event, which apparently excavated through the crust to some degree, was not energetic enough to shatter Vesta and destroy its basaltic crust. The fundamental reason that the large impact did not destroy Vesta’s crust is that it occurs in the gravity — rather than the strength — regime, as shown by detailed numerical simulations (Asphaug, 1997).

At the same time, collisions are argued to have produced the parent bodies of iron meteorites, widely interpreted to be the metallic cores of disrupted differentiated asteroids. These metallic cores were identified with the M-class asteroids; however, this interpretation has been challenged by the discovery of water of hydration on many M asteroids, leading to the proposed “W” (or wet) taxonomic class (Rivkin et al., 2000). However, the M-class asteroid 16 Psyche appears to be the collisionally exposed core of a parent body that was virtually identical to Vesta. How is it possible to disrupt the Psyche parent body and remove all traces of this giant collision (there is no family associated with Psyche today) except for Psyche itself, while still preserving the core of a differentiated body. The more general problem for collisional effects on asteroid materials is how to disrupt the dozens of parent bodies needed to explain the iron meteorite collection while eliminating from the asteroid belt most of the volumetrically more abundant dunite mantles of these bodies. Collisional grinding of this mantle material has been suggested, but such a process would need to be extremely efficient (see the section on the great dunite shortage in Bell et al. (1989); see also Burbine et al. (1996)).

Asteroid families are another observable that is the product of collisional evolution. There are approximately 25 reliable families and over 60 statistically significant clusters identified in asteroid proper elements [see Zappalà et al. (2002) and Bendjoya and Zappalà (2002) for detailed discussions of asteroid families].

Another constraint on the asteroid small size population comes from the cosmic-ray-exposure (CRE) ages of stony meteorites, which are measured to be around 20 m.y. The CRE age measures the time that the meteorite was in meter-sized bodies (or near the surface of a larger body); however, the collisional lifetime of meter-sized objects is calculated to be around 1 m.y. (O’Brien and Greenberg, 1999). Any successful asteroid collisional history must be consistent with the observed CRE ages.

Collisions have modified asteroid rotation rates over solar system history, so information about the collisional history of the belt is embedded in the spin rates of asteroids of all sizes. However, as discussed by Davis et al. (1989) and Farinella et al. (1992a), the uncertainties in modeling how collisions alter rotation rates are so large that any definitive work on this topic is yet to be done. However, recent progress has been made on understanding fragment rotation rates from catastrophic disruption events using numerical hydrocodes by Love and Ahrens (1997) and Asphaug and Scheeres (1999). These investigations provided deeper insights into fragmental spin rates as a function of fragment size as well as providing data on critical parameters needed to model the collisional evolution of asteroid spin rates.

In summary, any successful collisional evolution scenario for main-belt asteroids must satisfy the observable constraints imposed by the size distribution, the number of families, and the existence of an old basalt crust on Vesta; the CRE ages of strong meteorites; and the cratering projectile flux onto asteroids. Other observables, such as aster-

![Fig. 2. Nine estimates of the main-belt asteroid size distribution. The four models from Farinella et al. (1992b) are based on assuming N(>50 km) = 700, a cumulative diameter slope of –1.0 between 50 km > D > Dg and the Dohnanyi (1969) slope (cumulative diameter) slope of –2.5 for D < Dg. The values of Dg are taken to be 10 km, 25 km, 1 km, and 40 km, respectively. The model labeled “Galileo” is the adopted asteroid population for calculating the projectile flux onto Gaspra. They used the PLS slope (–1.95) for D > 0.175 km and the Dohnanyi slope for D < 0.175. The Davis et al. (1994) model derives from extrapolating the known population (assumed to be complete) at D = 44 km using the PLS slope and the Cellino et al. (1991) “corrected random” albedo model down to D = 5.5 km. The Durda et al. (1998) model is the fit to the asteroid size distribution determined by Jedicke and Metcalfe (1998), who used Spacewatch data to estimate the small size main-belt distribution. The SAM99 model is the Statistical Asteroid Model of Tedesco et al. (2002), who combined the observed size distribution of the main asteroid families with that of the background population into a single unified population estimate. The small size distribution is obtained by extrapolating the observed large size trends. Finally, the SDSS 2001 model is the fit to the Sloan Digital Sky Survey data by Ivezic et al. (2001), who find that a broken power law well represents their data. For 5 km < D < 40 km, the slope is ~3.0, while for 0.4 km < D < 5 km, a ~1.3 slope is a good fit to their data.](image-url)
Asteroid rotation rates, dust production, etc., are altered by, or are the product of, collisional processes. However, our understanding of how collisions affect them is not sufficiently mature at this time to allow imposition of additional constraints.

### 3. SCALING LAWS, BOUNDARY CONDITIONS, AND COLLISIONAL EVOLUTION

#### 3.1. Basic Collisional Parameters: Collision Rates and Impact Speeds

The frequency of collisions and the collisional impact speeds are fundamental quantities for collisional evolution studies. While the basic theory for calculating collision rates and speeds was developed by Öpik (1951) and Wetherill (1967) and applied by various workers in subsequent years, there has been recent work to refine such calculations and to extend them to both Trojans and Hildas. The intrinsic collision probability, $P_i$, gives the collision rate per unit cross-section area of target and projectile per unit time, with the actual number of collisions onto a target of radius $R_t$ by projectiles of radius $R_p$ within a time ($\Delta T$) is then given by

$$N_{coll} = \langle P_i \rangle (R_t + R_p)^2 V N_p \Delta T$$

where $V$ is the mean impact speed and $N_p$ is the spatial density of projectiles.

For each population several estimates of $\langle P_i \rangle$ and $V$ are given in Table 1. Different authors improved the Öpik/Wetherill analytical formulation with corrections related to specific features of the orbital distribution of the population under study or to statistical mechanics (Greenberg, 1982; Namiki and Binzel, 1991; Farinella and Davis, 1992; Bottke and Greenberg, 1993; Bottke et al., 1994; Vedder, 1996, 1998; Dell’Oro and Paolicchi, 1998; Dell’Oro et al., 1998, 2001). These methods also allow fast computations of updated values of $\langle P_i \rangle$ and $V$ whenever the discovery of new objects significantly enriches the known population. Alternatively, different authors have resorted to a direct numerical approach based on the integration of the orbits of the asteroids over a sufficiently long timespan. The derived dis-

<table>
<thead>
<tr>
<th>Asteroid Populations</th>
<th>Intrinsic Probability (10$^{-18}$ yr$^{-1}$ km$^{-2}$)</th>
<th>Impact Velocity (km s$^{-1}$)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Main Belt (MB)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MB-MB</td>
<td>2.85 ± 0.66</td>
<td>5.81 ± 1.88</td>
<td>Farinella and Davis (1992)</td>
</tr>
<tr>
<td>MB-MB</td>
<td>3.97</td>
<td>—</td>
<td>Yoshikawa and Nakamura (1994)</td>
</tr>
<tr>
<td>MB-MB</td>
<td>2.86</td>
<td>5.3</td>
<td>Bottke et al. (1994)</td>
</tr>
<tr>
<td>MB-SPC</td>
<td>1.51</td>
<td>10.86</td>
<td>Gil-Hutton (2000)</td>
</tr>
<tr>
<td>MB-UN</td>
<td>1.08</td>
<td>14.0</td>
<td>Gil-Hutton and Brunini (1999)</td>
</tr>
<tr>
<td><strong>Hildas (Hil)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hil-Hil</td>
<td>2.31 ± 0.10</td>
<td>3.09 ± 1.47</td>
<td>Dahlgren (1998)</td>
</tr>
<tr>
<td>Hil-Hil</td>
<td>1.93</td>
<td>3.36 ± 1.52</td>
<td>Dell’Oro et al. (2001)</td>
</tr>
<tr>
<td>Hil-Tro</td>
<td>0.24 ± 0.06</td>
<td>4.59 ± 1.71</td>
<td>Dahlgren (1998)</td>
</tr>
<tr>
<td>Hil-Tro</td>
<td>0.27</td>
<td>4.49 ± 1.68</td>
<td>Dell’Oro et al. (2001)</td>
</tr>
<tr>
<td>Hil-MB</td>
<td>0.62 ± 0.04</td>
<td>4.78 ± 1.78</td>
<td>Dahlgren (1998)</td>
</tr>
<tr>
<td>Hil-MB</td>
<td>0.66</td>
<td>4.80 ± 1.77</td>
<td>Dell’Oro et al. (2001)</td>
</tr>
<tr>
<td>Hil-UN</td>
<td>6.5</td>
<td>10.98</td>
<td>Gil-Hutton and Brunini (1999)</td>
</tr>
<tr>
<td><strong>Trojans (Tro)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>L4-L4</td>
<td>6.46 ± 0.09</td>
<td>4.90 ± 0.07</td>
<td>Marzari et al. (1996)</td>
</tr>
<tr>
<td>L5-L4</td>
<td>7.79 ± 0.67</td>
<td>4.66</td>
<td>Dell’Oro et al. (1998)</td>
</tr>
<tr>
<td>L5-L5</td>
<td>5.30 ± 0.10</td>
<td>4.89 ± 0.10</td>
<td>Marzari et al. (1996)</td>
</tr>
<tr>
<td>L5-L5</td>
<td>6.68 ± 0.18</td>
<td>4.51</td>
<td>Dell’Oro et al. (1998)</td>
</tr>
<tr>
<td>L4-SPC</td>
<td>0.33</td>
<td>6.78 ± 2.71</td>
<td>Dell’Oro et al. (2001)</td>
</tr>
<tr>
<td>L5-SPC</td>
<td>0.35</td>
<td>6.67 ± 2.59</td>
<td>Dell’Oro et al. (2001)</td>
</tr>
<tr>
<td>Tro-UN</td>
<td>0.50</td>
<td>8.19</td>
<td>Gil-Hutton and Brunini (1999)</td>
</tr>
</tbody>
</table>

For each population the values calculated by various authors with the relative errors, when available, are reported. The more recent values are usually derived from a larger sample of orbits and are probably more accurate. Collision rates with external populations are given for (a) short-period comets (SPC), and (b) Uranus/Neptune scattered planetesimals (UN).
tribution of close encounters and mutual speeds recorded during the integration can be extrapolated to infer the collision probability and characteristic impact speed (Yoshikawa and Nakamura, 1994; Marzari et al., 1996; Dahlgren, 1998).

It is interesting to note that \( P_i \) and \( V \) for each of the three populations treated here are rather similar, differing by at most a factor of 4 in \( P_i \) and a little more than a factor of 2 in \( V \). The collision rates between populations, though (e.g., Hildas with main-belt asteroids or Hildas with Trojans), can be an order of magnitude smaller than the intrapopulation collision rates. However, the collision speeds for the interpopulation collisions are similar to those within a population.

While various authors have attempted to infer the primordial population of asteroids by integrating backward in time from the present belt, the inherent instability of the collisional problem prevents such results from converging. Indeed, the earliest result of Dohnanyi (1969), which showed that under specific circumstances the asteroid population was collisionally relaxed and independent of the starting population, meant that it would be impossible to uniquely infer the initial asteroid population by going back in time from the present belt. All recent models of asteroid collisional evolution have assumed various hypothetical initial populations and propagated these forward in time using a variety of scaling laws, and have compared the terminal model population with the present observed belt.

### 3.2. Scaling Laws and Collisional Evolution

The modeling of the collisional evolution of populations of small bodies is generally performed in two main stages: (1) the simulation of the outcomes of single collisions, and (2) the integration of the equations of evolution, assuming given boundary conditions. A set of simplifying assumptions is usually made in these kind of models: Bodies are considered to be spherical and homogeneous, and they move and collide within a finite volume around the Sun with relative impact speeds determined by the orbits of the target and projectile. Simulations are performed by numerically integrating a given set of first-order, nonlinear, differential equations, taking into account the rate of variation of the population of asteroids of any given mass due to the number of fragments produced or destroyed by collisions, and to the removal of bodies by nongravitational effects (e.g., Campo Bagatán et al., 1994a).

A useful mathematical analysis for the collisional evolution of the asteroid belt was introduced some 30 years ago by Dohnanyi (1969, 1971) and independently by Hellyer (1970, 1971). Three crucial assumptions of these models are: (1) the interacting bodies are modeled as spheres of equal density, and their collisional cross-section is simply proportional to the squared sum of the radii; (2) all the collisional response parameters are size-independent, implying that fragmentation occurs for a fixed projectile-to-target mass ratio and no self-gravitational effect is taken into account; (3) the population has an upper cutoff in mass (a largest body, Ceres in the asteroid belt) but no lower cutoff. These workers analytically found a stationary solution to the governing differential equations for collisional evolution. To trace the detailed time history of collisional evolution and to include more realistic collisional physics requires numerical models. The first such models led to new insights regarding asteroid collisional history (Chapman and Davis, 1975; Davis et al., 1979) and numerical models have been used by nearly all workers subsequently.

The most important result from Dohnanyi’s work is that as the collisional process gives rise to a cascade of fragments shifting mass toward smaller and smaller sizes, a simple power-law equilibrium size distribution is approached, which has a cumulative diameter slope of \(-2.5\). Other authors have confirmed this result, both analytically and numerically (e.g., Paolicchi, 1994; Williams and Wetherill, 1994; Tanaka et al., 1996). In the real world, however, the basic assumptions of Dohnanyi’s theory are not fulfilled. It was realized early on that gravitational binding, a mass-dependent process, plays a dominant role in holding large asteroids together, hence the assumption that self-similar collisional outcomes could not hold in the gravity regime (Davis et al., 1979). Below we separately explore progress made in recent years in understanding the consequences of relaxing the above assumptions 2 (collisional parameters are size-independent) and 3 (no lower cutoff in the size distribution). Relaxing assumption 1 (on collisional cross-sections) has only a minor effect on the overall collisional evolution.

Recent collisional models replace Dohnanyi assumption 2 with a scaling law describing how the collisional outcomes vary with target and impactor size. In its most general form, a scaling law is simply an algorithm for extrapolating the outcomes of collisions for which we have direct, physical experimental experience to the outcomes of collisional events at size scales much larger or smaller than can be practically treated under laboratory conditions. One important component of scaling laws is determining the critical impact specific energy, \( Q_{D} \), the energy per unit target mass delivered by the projectile required for catastrophic disruption of the target, i.e., such that the largest resulting object has a mass one-half that of the original body. Notice that this object may be formed by partial reaccumulation of fragments due to self-gravity. Sometimes, as in laboratory experiments, the shattering impact specific energy is given in terms of the energy per unit target mass required for the catastrophic shattering of the target, \( Q_{S} \) (regardless of reaccumulation of fragments), such that the largest fragment produced has a mass of one-half that of the original target. The two definitions only differ in the regime in which the effect of gravity is much larger than that of solid-state forces.

A number of experimental and analytical studies have focused on quantifying just how \( Q_{D} \) scales with target size in the strength-dominated and gravity-dominated regimes (see Durda et al., 1998, for a recent review) and at what sizes the transition from strength-scaling to gravity-scaling takes place. Dimensional analyses (Farinella et al., 1982;
Housen and Holsapple, 1990), impact experiments (Holsapple, 1993; Housen and Holsapple, 1999), and hydrocode studies (Ryan, 1992; Benz and Asphaug, 1999) show that $Q_D^*$ for strength-dominated targets scales as roughly $D^{-0.24}$ to $D^{-0.61}$, where $D$ is the target diameter. On the other hand, determining the size dependence of $Q_D^*$ in the gravity-dominated regime has proven to be more difficult, particularly due to the lack of direct experimental experience in the disruption of asteroid-sized targets. Numerous analytical arguments and hydrocode studies (e.g., Davis et al., 1985; Housen and Holsapple, 1990; Love and Ahrens, 1996; Melosh and Ryan, 1997; Benz and Asphaug, 1999) yield size-dependent $Q_D^*$ that scale as $D^{1.13}$ to $D^{2}$. Recent hydrocode studies (Benz and Asphaug, 1999) and numerical collisional models (Durda et al., 1998) indicate that the transition from the strength-dominated to the gravity-dominated regime occurs at target diameters of about 100–300 m, while many earlier studies generally place the transition in the ~1–10-kilometer size range. Observations of asteroid spin rates (Pravec and Harris, 2000) confirm these theoretical findings, showing a lack of objects rotating with periods less than 2.2 h among asteroids with absolute magnitude $H < 22$, evidence that asteroids larger than a few hundred meters across are indeed mostly gravity-dominated aggregates with negligible tensile strength (“rubble piles”; see Richardson et al., 2002).

Durda (1993) and Durda and Dermott (1997) examined the influence of $Q_D^*$ on the shape of an evolving size distribution by showing that the power-law slope index of a population in collisional equilibrium is a function of the size dependence of $Q_D^*$. In the case of pure energy scaling (size-independent $Q_D^*$), the slope index of the size distribution is $\sim -2.5$, as Dohnanyi (1969) showed. When $Q_D^*$ decreases with increasing target size, as is the case in the strength-scaling regime, the slope index of the equilibrium size distribution is steeper than Dohnanyi’s; instead, when $Q_D^*$ increases, as in the gravity-scaling regime, the resulting slope index is shallower than $-2.5$.

Given the wide range of $Q_D^*$ proposed by different scaling laws, the question arises as to which one best describes the real world? Experiments involving asteroid-sized bodies are unfortunately as yet impossible to arrange, so the next best test is to compare model results using different scaling laws with asteroid collisional observables (section 2). Davis et al. (1994) tested several proposed scaling laws and found that strain-rate scaling of Housen et al. (1991) and simple energy scaling best satisfied all constraints. Further work testing scaling laws was carried out by Durda et al. (1998), who presented results from numerical experiments illustrating in a more systematic fashion the sensitivity of an evolved size distribution on the shape of the strength-scaling law. A linear relationship between $\log Q_D^*$ and $\log D$ yields a power-law size distribution; if there is nonlinearity in this relationship (due to a transition from strength- to gravity-scaling, for instance), then structure is introduced into the evolving size distribution due to a nonlinearity in collision lifetimes of the colliding objects.

Figure 3 shows three hypothetical scaling laws used to explore the behavior of evolved size distributions. For scaling law SL1a, at small sizes we assume a strength-scaling law with a $D^{-0.4}$ size dependence. At large sizes a gravity-scaling relationship with a $D^{1.1}$ size dependence is similar to the predictions of hydrocode models. Scaling law SL1b is similar to SL1a with the exception of a more gradual transition between the strength and gravity-scaling regimes. Scaling law SL2 is identical in shape to SL1a, but is 100x stronger everywhere.

Figure 4 shows the resulting evolved size distributions for colliding populations with the properties of disruption specific energy governed by scaling laws SL1a, SL1b, and SL2. In all three cases, the transition from strength- to gravity-scaling at $D = 100$ m results in a “bump” in the size distribution for objects just larger than this size. This bump results from the increased disruption lifetimes of objects in this size range over what they would have been had gravity-scaling not taken over to strengthen larger objects. Objects in this bump represent an excess supply of projectiles that can disrupt targets roughly an order of magnitude larger than themselves, resulting in a wavelike perturbation to the size distribution at larger sizes, as described in further detail below. The abruptness of the transition from strength-to gravity-scaling can significantly influence the shape of the resulting size distribution, with more gradual transitions (SL1b vs. SL1a) resulting in smaller amplitude waves. Greater strengths at all sizes (SL2 vs. SL1a) result in higher frequency waves, since the critical projectile size for target disruption is closer to the size of the targets themselves, and the mechanism driving the wave occurs over a smaller range of sizes.

These results demonstrate that evolved size distributions generated by collisional models depend strongly (and understandably) upon the shape of the size-strength scaling law. With this in mind, Durda et al. (1998) adopted a least-squares approach and adjusted the strength law for asteroi-
Dal bodies to obtain a best fit to the actual asteroid size distribution determined from the cataloged asteroids and Spacewatch data. Their derived strength scaling law, with a strength minimum at ~150 m diameter as independently suggested by hydrocode models (e.g., Benz and Asphaug, 1999) and observed rotation rates (e.g., Pravec and Harris, 2000), naturally gives rise to an evolved main-belt asteroid size distribution in good agreement with the two-bump structure observed in the actual size distribution of main-belt asteroids. Durda et al. (1998) interpret the bump in the size distribution between ~3 and 30 km (see Fig. 2) as a primary bump due to the transition from strength scaling to gravity scaling for asteroids larger than ~150 m; the well-known bump at ~50–200 km is a secondary feature resulting from a wavelike structure induced in the size distribution by the ~3–30-km primary bump.

### 3.3. Boundary Conditions and Collisional Evolution

Even more severe departures from simple power-law distributions result from relaxation of Dohnanyi assumption 3 (no lower cutoff in the size distribution). When a small-size cutoff in the initial distribution is imposed, a remarkable feature appearing in the simulations is that the final distribution is not a fixed-exponent power-law, but rather displays a wavy pattern of variable wavelength and amplitude, superimposed on an “average” power law of slope close to Dohnanyi’s canonical value.

The wavy structure is produced as follows. We divide the asteroid mass range into a series of discrete mass bins, so the smallest one in the distribution has particles removed only in two ways: collisions within the bin itself or by its particles hitting larger targets. Thus there is only a small number of bodies removed from this bin. The next larger bin, however, has bodies removed not only by the above processes, but also due to impacts by projectiles from the smallest bin. This leads to an enhanced removal rate from this bin, and its population is more rapidly depleted than that of the smallest bin. As we proceed to ever larger sizes, the depletion rate increases relative to that of smaller bins due to the ever-increasing number of projectiles per target body. Thus the population is increasingly depleted, and develops a steeper slope than Dohnanyi’s equilibrium value, up to the size bin for which the smallest available projectile can shatter its members. Beyond this size, larger target bins no longer have an enhanced depletion rate, but rather have a somewhat reduced removal rate due to the rapid decrease in the number of available projectiles with increasing size. This leads to a flattening of the size distribution and to a slope lower than the equilibrium one — thus a wavelike pattern arises. This pattern propagates itself to larger and larger sizes due to the rise and fall of the relative number of projectiles capable of shattering larger bodies. It is interesting to note that this shift propagates in such a way that the final distribution at sizes between 1 and 100 km is substantially changed when the cutoff is changed. Figure 5, for instance, shows the effect of shifting the cutoff size from 1 cm to 1 mm. Thus, the behavior of very small particles can in principle affect the size distribution of the observable asteroids in a very significant way. Only when averaged over several orders of magnitude in size does the overall size distribution still approximately match the canonical Dohnanyi equilibrium slope.

Is such a small-particle cutoff actually present in the actual asteroid population? The answer to this question is...
not obvious, but many nongravitational forces do indeed act on interplanetary matter and are efficient at different sizes: the solar wind, Poynting-Robertson drag, and the Yarkovsky effect. There is strong observational evidence that most of the micrometeoroid mass is concentrated at particles sizes around 100 µm, and that the corresponding mass distribution flattens considerably for smaller sizes (Grün et al., 1985; Love and Brownlee, 1993). Thus the real low-mass cutoff of the asteroid population probably lies at sizes of micrometers or smaller. Solar radiation pressure on particles of size ~1 µm (Burns et al., 1979) converts them into hyperbolic β meteoroids, which, according to Grün et al. (1985), provide the main loss mechanism for the meteoritic complex. Durda (1993) and Durda and Dermott (1997) imposed a small-mass cutoff in their collisional model that empirically matched the Grün et al. (1985) data and found it too gradual to induce significant wave structure in the evolved population. Models that attempt to simulate the removal of small particles from the colliding population through radiation forces (Campo Bagatian et al., 1994a) do show a noticeable wave structure only when Poynting Robertson drag is artificially enhanced; the Yarkovsky effect — much more effective than the Poynting-Robertson — might instead be responsible for such a pattern. Clearly, more work is needed in this area to explicitly treat within collisional models the various nongravitational forces that act to remove small objects from the asteroid population and to determine how strong and abrupt a small-mass cutoff must be in order for wavelike features to appear in the evolved size distribution.

3.4. Model Parameters and Collisional Evolution

Besides the scaling laws for $Q_D$ or $Q_S$, which affect the disruption/fragmentation threshold and the number of fragments produced in any collision, fragmentation models depend on a number of poorly known critical parameters that govern the mass distribution and the reaccumulation of fragments after their formation (see Campo Bagatian et al., 1994b). Numerical models generally calculate the fragmental size distribution, assumed to be a power law, from $Q_S^*$ and the collisional energy. The inelasticity parameter, $f_{KE}$, which determines what fraction of the collisional kinetic energy goes into the fragments, is used to calculate the ejecta fraction that escapes the bodies’ gravity. Its value depends on the composition, internal structure, and size of the bodies, and it is often taken between 1% and 10%. On the other hand, $f_{KE}$ is implicitly embedded in $Q_D$ in that a large $Q_D$ implies a small $f_{KE}$, while a small $Q_D$ goes with a larger $f_{KE}$ (Campo Bagatian et al., 2001).

The fragment mass-speed relation is critical to calculating collisional outcomes and relationship of the form $V(m) = Cm^{-r}$ (Nakamura et al., 1992; Giblin, 1998), with $r$ ranging from 0 (no mass-velocity dependence) to 1/6, as found by Giblin (1998). The effects of this dependence have been studied by Petit and Farinella (1993) and Campo Bagatian et al. (1994b). Even a shallow dependence of ejection speed on fragment mass makes a significant difference in the amount of mass that can be reaccumulated by objects in the range 1 to 100 km (Campo Bagatian et al., 2001). Also, the size at which a significant ratio of aggregate rubble-pile asteroids is present seems to depend strongly on the reaccumulation model. For example, assuming a mass-velocity relationship with $r = 1/6$ provides some 10% rubble-pile asteroids for 2-km-sized bodies, while $r = 0$ prevents reaccumulated objects for sizes below some 10 km (Fig. 6). It is thus clear that further efforts are necessary to understand the dependence of ejection speeds of fragments on their masses. On the other hand, recent work using hydrocodes to model collisional fragmentation (Benz and Asphaug, 1999; Michel et al., 2001; see also section 5) suggests that essentially all asteroids from ~1 to 100 km diameter are reaccumulated rubble piles.

![Fig. 5.](image) Propagation of the wave effect with different small-size cutoffs (solid line: 1 cm; dashed line: 1 mm).

![Fig. 6.](image) Fraction of rubble piles in the asteroid belt as a function of size, according to different reaccumulation models (solid line: mass-velocity dependence with $r = 1/6$; dashed line: no mass-velocity relation, i.e., $r = 0$), and keeping $f_{KE} = 0.01$ (see text).
Recent models of asteroid collisional evolution consider two different kinds of objects (Campo Bagiati et al., 2001): monoliths (high $f_{KE}$), and rubble-pile reaccumulated objects (low $f_{KE}$), usually called rubble-piles, or aggregate asteroids. The number of aggregate asteroids present in the main asteroid belt can be estimated through collisional modeling to range from 30% to 100% in the range 10 to 200 km, depending on the adopted scaling law, the exponent r, and the choice of other collisional parameters, especially the inelasticity parameter, $f_{KE}$.

A characteristic pattern of collisional evolution outcomes is that the final size distribution of bodies appears to be fairly insensitive to the initial conditions, provided most of the mass is in bodies that can be collisionally disrupted at impact speeds characterizing the population. For any population characterized by an impact speed, $v$, there is a largest-sized body that can be collisionally disrupted even if hit by an equally large body. As long as most of the population mass is in bodies smaller than this critical size, the final distribution is essentially independent of the starting distribution. Even for the extreme initial distribution without bodies smaller than 1 km, the subkilometer size range is populated by fragments within a few million years, and the stationary distribution is attained.

3.5. Asteroid Observables and Collisional Evolution

The previous three sections tested asteroid collisional evolution models against the observed size distribution of asteroids, but as noted in section 2, there are other collisionally produced observables:

Families. Marzari et al. (1999) used the ~85 recognized families and statistically significant grouping recognized in the present belt as a constraint on the overall asteroid collisional history. Their work showed that a small mass initial belt best reproduced the observed number and types of families produced by disruption of parent bodies larger than 100 km diameter.

Vesta’s crust. Preservation of the basaltic crust of Vesta requires that the mass of the asteroid belt was at most several times the mass of the present belt at the time that the present dynamical environment was established. This statement assumes a power-law size distribution for the early asteroid population; it would be possible for the early asteroid belt to contain significantly more mass than it does today if such mass were in a few massive bodies that were removed from the belt by dynamical, rather than collisional, processes. Work by Davis et al. (1984) and Davis et al. (1994) showed the range of initial asteroid populations that were consistent with the preservation of Vesta’s crust.

Cosmic-ray exposure ages of meteorites and the Gaspra/Ida cratering record. Reconciling these constraints with the overall asteroid size distribution remains a challenging problem (O’Brien and Greenberg, 2001). Collisional models that include Poynting-Robertson drag and the Yarkovsky effect, along with hydrocode-derived scaling laws, offer promise for understanding the asteroid size distribution from the largest bodies down to subdecameter-sized asteroids.

4. COLLISIONAL EVOLUTION OF TROJANS AND HILDAS

4.1. Trojans

Jovian Trojan asteroids have a peculiar dynamical behavior: They librate around the Lagrangian equilibrium points L4 and L5 located at the maxima of the “pseudo-potential” defined within the frame of the restricted three-body problem (Greenberg and Davis, 1978). The resonant link with Jupiter confines them in a limited volume of space and, notwithstanding that they have an average high inclination on the ecliptic, their collisional activity is comparable to that of main-belt asteroids. Collisions for Trojans are a relevant evolutionary mechanism not only because they grind down the primordial size distribution, but also because they significantly alter the parameters of the librational motion. The velocity imparted to fragments in a collision generate new orbital parameters that, for librating bodies, may even drive them out of the resonance. The present Trojan population has almost surely lost memory of the primordial distribution of libration amplitude and proper eccentricity.

The first step to quantitatively model the collisional evolution of Trojans is to derive precise estimates of the typical collisional frequencies and impact velocities for the two swarms. Since the librational motion sets kinematical constraints to the values of the orbital angles respective to those of Jupiter, the standard statistical approaches based on Wetherill’s (1967) theory, even in more recent formulations (Bottke and Greenberg, 1993; Vedder, 1996, 1998), cannot be applied. Critical assumption of these models is that the perihelion argument and longitude of nodes of the bodies precess regularly, while this condition is violated for Trojans. More sophisticated analytical models like that of Dell’Oro et al. (1998) or, alternatively, a direct numerical approach (Marzari et al., 1996) have been employed to determine the collisional parameters (Table 1).

Armed with improved values of the impact parameters, Marzari et al. (1997) numerically simulated the collisional evolution of Trojan asteroids over the age of the solar system. They found that the slope of a primordial planetesimal-like initial population truncated below 10 km in diameter and with a steep slope (~5.5 incremental) gives a good match to the present population (Fig. 7a). An initial distribution of this kind would have been collisionally very active since it is far from Dohnanyi’s equilibrium slope and, as a consequence, we would expect a large number of asteroidal families to be produced over solar system history. Some of these families have been identified, even with a sample of only 174 Trojan orbits by Milani (1993, 1994), and by Beaugé and Roig (2001), based on a set of 533 Trojans. Another important aspect of the collisional evolution is the large flux of breakup fragments out of the Trojan swarms. As pointed
out by Marzari et al. (1995), the collisional disruption of a Trojan asteroid can inject some fragments into unstable orbits that end up in Jupiter-crossing chaotic cometary orbits. A quantitative estimate of the actual flux based on the numerical simulation of Marzari et al. (1997) suggests that the Trojan swarms could supply approximately 10% of the short-period comet and Centaurs populations. This dynamical connection between Trojans and short-period comets is reinforced by the spectrophotometric similarity with D-type asteroids, which are dominant among Trojans, and inactive cometary nuclei (Hartmann et al., 1987; Shoemaker et al., 1989).

The primordial Trojan population found by Marzari et al. (1997) showed only a possible lower limit to the real primordial population due to neglecting the depleting effect of dynamics on the population. As shown by Levison et al. (1997), the stability region populated by present Trojans in the phase space is surrounded by isochrone curves of limited lifetime. The primordial population of Trojans had then two independent sink mechanisms: collisions and long-term dynamical diffusion. Collisions grind down larger bodies to small fragments; dynamical diffusion ejects them out of the swarm into chaotic orbits. However, these mechanisms also work in synergy: Collisions move the bodies in the phase space, possibly refilling regions cleared up by instability, and at the same time, bodies on unstable orbits take part in the collisional evolution over a limited timespan before escaping. Further work is needed to understand the long-term effects on the size and orbital distribution of the Trojan population.

4.2. Hildas

The Hildas are objects that reside in the 3:2 mean-motion resonance with Jupiter at ~4.0 AU. These asteroids are relatively isolated from physical interactions with objects of the main belt, and are weakly coupled to other outer-belt asteroids such as Cybeles or Trojans. This is the reason their collisional activity is not very intense: The intrinsic collision probability for Hilda asteroids is lower than for main-belt Cybele and Trojan asteroids by about a factor of 2.2 and 3, respectively (Dahlgren, 1998; Dell’Oro et al., 2001).

The low collisional activity is the result of two effects. First, the 3:2 mean motion resonance with Jupiter has a 0.1-AU-wide dynamically stable zone, surrounded by a strongly chaotic boundary with very short characteristic diffusion times (Ferraz-Mello et al., 1998). Therefore, if any asteroid is extracted from the resonance, it is quickly ejected far away from the stable zone, no longer participating in the collisional evolution of the population. Then, if a fragment produced by a collision reaches a relative velocity enough to escape from the resonance, a depletion of the size distribution occurs (Gil-Hutton and Brunini, 2000). Assuming that the relative velocity of a fragment is small with respect.
to the orbital velocity of the parent body \( V_c \), the semimajor axis change is

\[
\Delta a = 2a \frac{\Delta V_r}{V_c}
\]

which is a zeroth-order form of Gauss’ equation, where \( a \) is the semimajor axis and \( \Delta V_r \) is the component of the relative velocity along the direction of the motion. Using \( a = 3.96 \) AU, and \( \Delta a \approx 0.05 \) AU, and assuming that the relative velocity of the fragment with respect to the parent body is equally partitioned between the radial, tangential, and normal components, the ejection velocity needed to escape from the center of the resonance is \( \Delta V > 0.163 \) km/s. Thus, small fragments produced by collisions closer to the boundary of the resonance or having large relative velocities can easily escape, depleting the Hildas’ size distribution and producing a permanent loss of projectiles.

Second, the initial mass for the Hildas was not large. It is possible to obtain a rough value for it using the model of planetary nebulae mass distribution proposed by Weidenschilling (1977) to calculate the ratio between the total mass of the main-belt and Hildas region. Another possibility is to assume that the number of objects currently observed is proportional to the initial mass, and find the ratio between the number of objects larger than a certain radius that are observed in both populations. These methods give similar results: The initial mass for the Hildas was \( \approx 22 \times \) less than the initial mass in the main belt [between \( 1.5 \times 10^{-3} \) and \( 3.6 \times 10^{-5} M_{\oplus} \) (Gil-Hutton and Brunini, 2000)].

As a consequence of these effects, it is expected that there has been only a modest collisional attrition to the primordial size distribution. Nevertheless, the observed size distribution shows a change in the slope at a radius of \( \approx 12 \) km, which was traditionally explained as an observational bias due to the faintness of small objects in the outer belt. However, it could be the result of a collisional process: Fernández and Ip (1983) and Brunini and Fernández (1999) found that the accretion process of Uranus and Neptune was extremely inefficient and a large number of bodies was scattered into the inner solar system in a period not longer than few times \( 10^7 \) yr (see section 4.3). Using this result, Gil-Hutton and Brunini (2000) found that collisions between Hildas and scattered planetesimals from the Uranus-Neptune zone could produce a large number of fragments that could easily escape from the resonance. Thus, the current size distribution of Hildas could be a result of an intense collisional process, which preferentially depletes the small end of the population, leaving the Hildas without enough objects to have further collisional evolution. As a consequence of the fast depletion and low collisional activity after the planetesimal bombardment, it is possible that the population of currently observed Hildas is complete or nearly complete at sizes \( \geq 20 \) km in diameter (Fig. 7b), and the slope change in the size distribution could be the result of this collisional process (Gil-Hutton and Brunini, 2000). This result is contrary to the usual assumption that the outer belt observed populations are far from complete, but it is observationally testable: Any future survey of Hildas asteroids should not modify the currently observed size distribution to a substantial degree.

4.3. Interrelations with Other Populations

The existence of planetesimal populations is an integral part of planetary formation by accretion. Once the proto-planets grew to masses several tenths the current ones, they started to scatter unaccreted bodies well beyond their accretion zones. As a possible byproduct of this process, large numbers of bodies from these populations diffused inward, reaching the inner solar system, and a certain fraction of them have eventually interacted with asteroids in the main belt.

Collisional interactions between the asteroid belt and scattered planetesimals from the Uranus and Neptune zone were first considered by Brunini and Gil-Hutton (1998). The possible presence of a nonnegligible amount of H and He in the outer planetary nebula (Pollack et al., 1996) suggests that the formation timescales of Uranus and Neptune could not have been much longer than the timescale of dissipation of the gaseous component of the nebula, perhaps not longer than a few \( 10^7 \) yr (Brunini and Fernández, 1999). A significant number of scattered planetesimals reached the region of the inner planets (Fernández and Ip, 1983), some of which collided with asteroids in the main belt. Gil-Hutton and Brunini (1999) studied the collisional process between these populations and found that the planetesimal bombardment was extremely efficient, removing mass of the primordial asteroid belt in a very short time, allowing very massive initial populations. In spite of the fact that the collisional process was very intense (a total mass infall of \( 15 M_{\oplus} \) was scattered to the inner solar system from the Uranus-Neptune zone), it lasted only \( 1 \times 3 \times 10^7 \) yr, so Vesta had a good chance of retaining its basaltic crust largely intact in this scenario.

The interaction between the asteroid belt and the short-period comet population was considered by Gil-Hutton (2000). This author obtained values for the intrinsic collision probabilities and collision velocities of 159 short-period comets (SPC) colliding with the main belt using the algorithm developed by Bottke et al. (1994; see Table 1). The intrinsic collision probability is about one-half that of main-belt collisions, but the mean impact speed is significantly higher. However, SPC are not a serious threat for asteroids due to their small population.

5. FUTURE CHALLENGES

Asteroid collision studies made sizeable strides in the past decade, particularly with the increased volume of data on the asteroid size distribution and a better understanding of fragmentation physics. However, there is still a wide discrepancy in scaling-law prediction for the collisional energy needed to disrupt bodies. As numerical models become more sophisticated and computers ever faster, new insights are expected on this fundamental aspect of collisional studies.
Recent results have been presented based on the marriage of smooth particle hydrocode calculations of the fragmentation process with an N-body integrator to follow the trajectories of fragments once the material interactions have ceased (Michel et al., 2001; Durda et al., 2001). Michel et al. (2001) have successfully reproduced the observed size distribution for asteroid families with such an approach and, at the same time, have introduced a new paradigm for collisional outcomes. In contrast to the earlier work, which argued that the size distribution from disruptive collisions was controlled by shattering the material bonds of the body, the new model suggests that gravitational reaccumulation is the fundamental mechanism for forming individual bodies during disruptive collisions. Hence, the velocity field established by the collision would determine the number and size of fragments, not the propagation of cracks that fracture material bonds. In the Michel et al. (2001) model, the collision is sufficiently energetic to pulverize the material of the body to very small sizes, but many of these fragments gravitationally reaccumulate to form the fragments that make up asteroid families. Further work is certainly needed in this area, but this approach suggests a very different physical basis for understanding the outcomes of disruptive collisions.

Variably all asteroid collisional studies point to a small mass initial belt, i.e., the total mass of the asteroid belt was only a few times larger than the present belt at the time the present disruptive dynamical environment was established. What is needed now is to link collisional models with accretion models for the formation of asteroids to produce a seamless, self-consistent scenario for the origin and evolution of asteroids. Work by Wetherill (1992) and Chambers and Wetherill (2001) suggests that massive bodies accreted in the primordial asteroid belt. These bodies acted to gravitationally stir up the population, thus terminating accretion among small bodies. The massive bodies were then removed by mutual gravitational scattering into resonances in the asteroid belt that formed when Jupiter grew to be a massive body. The residual mass from this process is what we see as the asteroid belt today.

Modeling the collisional evolution of asteroids populations is still a very complex problem, due to the fact that a number of poorly known free parameters are embedded in the modeling of the involved physical phenomena. More observational data and more refined modeling techniques are needed in order to get a comprehensive and self-consistent picture of fragmentation and collisional evolution of small bodies populations. Specific challenges in this field for the next decade include (1) achieving a better understanding, and accurate modeling, of the collisional response of rubble-pile asteroids to energetic impacts; (2) developing a numerical model that combines collisional processes and dynamics, e.g., the collisional evolution of asteroids including the Poynting-Robertson and Yarkovsky effects, the major dynamical resonances and, possibly, even three-body resonances; and (3) modeling the collisional evolution of asteroid rotations.

The relatively isolated Hildas and Trojans regions are natural laboratories to compare their collisional evolution with that of main-belt asteroids. As in the case for main-belt asteroids, one needs better knowledge about the size distribution of these populations down to kilometer-sized bodies. Also, combining dynamical depletion mechanism with collisional models will yield deeper insights into the evolution of these populations. The next decade promises to show major progress in deciphering the collisional history of small body populations.

REFERENCES


