Irregular Satellites of the Giant Planets

Philip D. Nicholson
Cornell University

Matija Ćuk
University of British Columbia

Scott S. Sheppard
Carnegie Institution of Washington

David Nesvorný
Southwest Research Institute

Torrence V. Johnson
Jet Propulsion Laboratory

The irregular satellites of the outer planets, whose population now numbers over 100, are likely to have been captured from heliocentric orbit during the early period of solar system history. They may thus constitute an intact sample of the planetesimals that accreted to form the cores of the jovian planets. Ranging in diameter from ~2 km to over 300 km, these bodies overlap the lower end of the presently known population of transneptunian objects (TNOs). Their size distributions, however, appear to be significantly shallower than that of TNOs of comparable size, suggesting either collisional evolution or a size-dependent capture probability. Several tight orbital groupings at Jupiter, supported by similarities in color, attest to a common origin followed by collisional disruption, akin to that of asteroid families. But with the limited data available to date, this does not appear to be the case at Uranus or Neptune, while the situation at Saturn is unclear. Very limited spectral evidence suggests an origin of the jovian irregulars in the outer asteroid belt, but Saturn’s Phoebe and Neptune’s Nereid have surfaces dominated by water ice, suggesting an outer solar system origin. The short-term dynamics of many of the irregular satellites are dominated by large-amplitude coupled oscillations in eccentricity and inclination and offer several novel features, including secular resonances. Overall, the orbital distributions of the irregulars seem to be controlled by their long-term stability against solar and planetary perturbations. The details of the process(es) whereby the irregular satellites were captured remain enigmatic, despite significant progress in recent years. Earlier ideas of accidental disruptive collisions within Jupiter’s Hill sphere or aerodynamic capture within a circumplanetary nebula have been found wanting and have largely given way to more exotic theories involving planetary migration and/or close encounters between the outer planets. With the Cassini flyby of Phoebe in June 2004, which revealed a complex, volatile-rich surface and a bulk density similar to that of Pluto, we may have had our first closeup look at an average-sized Kuiper belt object.

1. INTRODUCTION

Planetary satellites are conventionally divided into two major classes, based on their orbital characteristics and presumed origins (Burns, 1986a; Stevenson et al., 1986; Peale, 1999). The inner, regular satellites move on short-period, near-circular orbits in or very close to their parent planets’ equatorial planes, and are generally believed to have formed in situ via accretion from a protoplanetary nebula [see, e.g., Canup and Ward (2002) and Mosqueira and Estrada (2003) for recent models]. By contrast, the outer irregular satellites move on orbits with periods on the order of 1–10 yr, which are frequently highly eccentric and inclined. The latter bodies are generally believed to have been captured from heliocentric orbit during the final phases of planetary accretion (Pollack et al., 1979; Heppenheimer and Porco, 1977) or via a collisional mechanism at some later time (Colombo and Franklin, 1971), but there is no consensus on a single model. Recent years have seen an explosion in the number of known outer satellites of the jovian planets, due to deep imaging searches carried out with wide-field CCD cameras on large-aperture telescopes (Gladman et al., 1998a, 2000, 2001a; Sheppard and Jewitt, 2003; Holman et al., 2004; Kavelaars et al., 2004; Sheppard et al.,...
The advent of photographic plates in the late nineteenth century led to a burst of satellite discoveries, in a field that had lain dormant since the visual discoveries of Hyperion, Ariel, Umbriel, and Triton by W. Lassell and W. Bond in 1846–1851 and Phobos and Deimos by A. Hall in 1877. Almost all the faint, newly found objects proved to be what came to be known as irregular satellites: first Saturn’s Phoebe ($m_V = 16.5$), discovered by W. Pickering at Harvard’s southern station in 1898, followed over the next 16 years by Jupiter’s Himalia ($m_V = 14.6$), Elara (16.3), Paphsae (17.0), and Sinope (18.1) discovered by C. Perrine, P. Melotte, and S. Nicholson. Working at Mount Wilson, S. Nicholson discovered three additional 18th-mag jovian irregulars between 1938 and 1951: Lysithea, Carme, and Ananke. In the 1940s and 1950s, G. Kuiper undertook photographic surveys for new satellites at Uranus and Neptune using the McDonald 82-inch telescope, resulting in the discovery of Nereid in 1949. All these surveys depended on manual comparison of pairs of photographic plates, generally using blink comparators to superimpose the two images. The final discovery of the photographic era was Leda (J XIII), discovered by Kowal et al. (1975) on plates taken with the 48-inch Schmidt telescope at Palomar. At mean opposition magnitudes of 19.5 and 19.7, Leda and Nereid represented the practical limit for the detection of moving objects on long-exposure plates.

Much of this early survey work is summarized by Kuiper (1961), including his own extensive investigations, while some later but unsuccessful searches are reviewed by Burns (1986a). In a postscript to the era of photographic discoveries, we note the report by Kowal et al. of a possible ninth irregular jovian satellite, later designated S/1975 J1 (see IAU Circular 2855). This object remained unconfirmed for a quarter century until it was unexpectedly and independently rediscovered by Holman et al. and Sheppard et al. in CCD images in 2000; it was subsequently named J XVIII Themisto (see IAU Circular 7525).

2.2. CCD Surveys

Photographic plates are very inefficient for detecting moving objects. They do not reach faint magnitudes on short timescales, are very time intensive to handle, and are hard to compare in order to find moving objects. Additional irregular satellite discoveries would require more sensitive detectors that covered large fields of view (Fig. 1). In the mid 1990s digital CCDs finally were able to cover significant areas of the sky per exposure (10–30 arcmin on a side compared to only a few arcminutes previously). Since 1997, with the discovery of Caliban and Sycorax at Uranus by Gladman et al. (1998a), 95 irregular satellites have been discovered around the giant planets (Gladman et al., 2000, 2001a; Sheppard and Jewitt, 2003; Holman et al., 2004; Kavelaars et al., 2004; Sheppard et al., 2005, 2006). With these new discoveries Jupiter’s known irregular satellite population has grown from 8 to 55, Saturn’s from 1 to 35, Uranus’ from 0 to 9, and Neptune’s from 2 (if we include Triton) to 7.

Table 1 provides some basic information about the current state of the surveys for irregular satellites around the giant planets. The known irregular satellites (106 as of October 2006) now greatly outnumber the known regular satellites (Fig. 1). Because of its proximity to the Earth, Jupiter currently has the largest irregular satellite population but it is expected that each giant planet’s environs may contain similar numbers of such objects (Jewitt and Sheppard, 2005).

![Fig. 1. The number of irregular and regular satellites discovered since the late 1800s. Several key technological advances that resulted in a jump in the pace of discoveries are identified.](image-url)
CCDs are not only more sensitive than photographic plates but they allow for the use of computers, which can process the vast amount of data generated with relative ease. Two main techniques have been used with the images from CCDs to discover irregular satellites. Both techniques require observations of the planets to be made when they are near opposition. This ensures that the dominant apparent motion is parallactic, allowing the distance to the object to be calculated. At opposition the movement of an object is simply related to its heliocentric distance, r, by (Luu and Jewitt, 1988)

\[
\frac{d\Omega}{dt} = 148 \left( \frac{1 - r^{-0.5}}{r - 1} \right)
\]

where r is expressed in AU and \(d\Omega/dt\) is the rate of change of apparent position of the object relative to the stellar background in arcsec hr\(^{-1}\). This allows foreground main-belt asteroids and background Kuiper belt objects (KBOs) to be distinguished easily from possible outer planetary satellites, which have motions typically within a few arcseconds per day of their host planet.

The first technique requires the observer to take three images of a region of sky near the planet spread out over a few hours. Each set of three images per field is then searched using computer programs that detect objects moving relative to the background stars and galaxies near the rate determined from equation (1). Once flagged, these objects are followed up during the next few weeks and months to confirm their satellite status and to obtain preliminary orbits.

The second technique, first used for deep transneptunian object (TNO) searches and described by Gladman et al. (1999b) and Petit et al. (2004), is more computer-intensive and can only be used efficiently to search for satellites of Uranus and Neptune since these satellites have apparent motions almost exactly the same as those of their host planets (Holman et al., 2004; Kavelaars et al., 2004). In this technique several tens of images are taken of the same field over a period of a few hours. These images are then shifted at the planet’s rate and combined using a median filter. Satellites will show up as point sources while foreground and background objects will disappear or be smeared out by the shifting and combining. This technique prevents the CCDs from saturating on the sky while increasing the signal-to-noise of any satellite detection. fainter satellites can be found with this method, but as it is very telescope-time-intensive, only limited areas of sky can be searched on each night.

Unsuccessful searches for irregular satellites of Mars have been carried out by Showalter et al. (2001) and Sheppard et al. (2004), while Stern et al. (1991, 1994) and Nicholson and Gladman (2006) have reported similar negative searches for irregular satellites at Pluto. [Two newly discovered satellites of Pluto, II Nix and III Hydra (Weaver et al., 2006) both orbit close to the planet’s equatorial plane and are best classified as regular satellites.]

2.3. Orbital Groupings

Even a cursory inspection of the mean orbital elements of the irregular satellites of Jupiter and Saturn reveals several obvious groupings, illustrated in Plate 8. Prior to 1999, Jupiter had four known prograde satellites, dominated by 160-km-diameter Himalia and tightly clustered in a, e, and i, plus a much looser grouping of four retrograde satellites centered on Pasiphae (D = 60 km) and Carme (45 km). These were generally considered to represent fragments from the disruption — perhaps incidental to their capture — of two parent objects (Burns, 1986b; Colombo and Franklin, 1971).

With the current census of 55 irregular jovian satellites, at least 3 and perhaps as many as 4 distinct groups can be identified in a–e–i space (Nesvorný et al., 2003; Sheppard and Jewitt, 2003): one prograde and two to three retrograde. The original prograde group with a\(_e\) = 11.5 and i = 28° remains, consisting of Himalia, Elara, Lysitea and Leda, possibly augmented by S/2000 J11, although this object has not been recovered and is now considered “lost.” (We use the shorthand notation \(a_e = a/10^6\) km.) In addition, there are two isolated high-inclination objects: Themisto (\(a_e = 7.5, i = 43°\)) and Carpo (\(a_e = 17.0, i = 51°\)). All remaining jovians are retrograde, and were divided by Sheppard and Jewitt (2003) into three groups centered on Ananke (\(a_e = 21, i = 149°\), Pasiphae (\(a_e = 24, i = 151°\)), and Carme (\(a_e = 23, i = 165°\)). With 24 additional satellites discovered in 2002 and 2003, the Ananke and Carme groups have become more sharply delineated, but the Pasiphae group remains rather diffuse. The Ananke group has at least 8 smaller named members (Euanthe, Harpalyke, Hermippe, Iocaste, Mneme, Praxidike, Thelxinoe, and Thyon), while Carme has at least another 11 (Aitne, Arche, Chaldene, Erinome, Eukelade, Isonoe, Kale, Kallichore, Kalyke, Pasithee, and Taygete). Euporie and Orthosie, grouped with Ananke by Sheppard and Jewitt (2003), now appear more likely to be outliers (Nesvorný et al., 2003). The Pasiphae group may contain as many as 10 other named objects, including 40-km-diameter Sinope, Autonoe, Cyllene, Eurydome, Megale, and Sponde (Sheppard et al., 2003), but its exact membership is unclear. Questionable members are Helike (whose latest orbit puts it nearer the Ananke group), Hegemone, Aode, and Callirrhoe.

The saturnian irregulars comprise two relatively tight prograde groups in a–e–i space: Albiorix, Erriapo, Tarvos, and S/2004 S11 at \(a_e = 17, i = 34°\) and the Kiviuq/Ijiraq pair at \(a_e = 11.4, i = 46°\). Paaliaq and Siarnaq share the latter group’s inclination, but are at much larger semimajor axes of \(a_e = 15\) and 18, respectively. Excepting Skathi (S/2000 S8), all the retrograde saturnian satellites were lumped together by Gladman et al. (2001a) into a rather loose Phoebe inclination group with \(i = 170°\). Unlike the situation at Jupiter, the saturnian retrograde satellites do not form well-defined groups in a–e–i space. Instead, we may group them roughly by inclination alone. Including a dozen new satellites reported by Jewitt et al. in IAU Circular 8523, a
The Solar System Beyond Neptune

smaller, tighter Phoebe group has emerged at $i = 175^\circ$ (with Suttungr, Thrymr, Ymir, and S/2004 S8), while Mundilfari at $i = 168^\circ$ is joined by S/2004 S7, S10, S12, S13, S14, S16, and S17. Skathi at $i = 153^\circ$ has acquired four comrades (Narvi = S/2003 S1, S/2004 S9, S15 and S18). These retrograde saturnian inclination groups encompass a wide range of semimajor axes and eccentricities, raising considerable doubt as to whether they represent actual “families” in the genetic sense (see section 5.3). Nine new retrograde saturnian satellites reported by Sheppard et al. (see IAU Circular 8727) all appear to fall within the above three inclination groups.

At Uranus the situation is rather different. First, all but one (Margaret) of the nine known irregulars are retrograde. Second, both Kavelaars et al. (2004) and Sheppard et al. (2005) concluded that while the inclination distribution at Uranus is essentially random, with the possible exception of (2005) concluded that while the inclination distribution at Uranus is essentially random, with the possible exception of the pair Caliban and Stephano at $a_6 = 7.5$, $i = 142^\circ$, there is a statistically-significant separation into two groups in $a$–$e$ space. The inner, more circular group at $a_6 = 7$, $e = 0.2$ consists of Caliban, Stephano, Francisco, and Trinculo, while the outer, more eccentric group at $a_6 = 16$, $e = 0.5$ consists of Sycorax, Setebos, Prospero, and Ferdinand.

With only seven irregular satellites currently known (five if Triton and Nereid are excluded as atypically large and possibly sharing a unique dynamical history, as discussed in section 6.2), the classification of Neptune’s outer satellites remains problematic (Holman et al., 2004; Sheppard et al., 2006). The three known prograde satellites, Nereid, Laomedea = S/2002 N3, and Sao = S/2002 N2, have very different inclinations of $i = 7^\circ$, 35$^\circ$, and 48$^\circ$ respectively. Three of the four retrograde satellites (Halimede = S/2002 N1, Neso = S/2002 N4, and Psamathe = S/2003 N1) form a fairly tight grouping in inclination at $i \sim 135^\circ$, while the latter pair also have quite similar values of $a$ and $e$.

As discussed further in section 3.1, the rapid decrease in brightness of satellites with increasing heliocentric distance means that current surveys at Uranus and Neptune are almost certainly much less complete than those at Jupiter and Saturn, raising the distinct possibility that we are seeing only the largest member or two of each orbital group. It is worth recalling that what appeared a decade ago to be a single loose cluster of four retrograde jovian irregulars is now seen as three distinct and much tighter groups.

2.4. Nomenclature

Historically, different conventions have been followed in naming the irregular satellites of the various planets. The outer satellites of Jupiter have names derived from classical Roman or Greek mythology and are lovers or descendents of Jupiter or Zeus. The jovian prograde satellites with inclinations near 28$^\circ$ and presumably related to Himalia have names ending in “$a$”; more highly inclined prograde objects have names ending in “$o$”; and retrograde objects have names ending in “$e$.” At Saturn, where the inner satellites are named after Titans, the outer irregulars are named after giants from Gallic ($i = 35^\circ$), Inuit ($i = 45^\circ$), and Norse (retrograde) mythology, Phoebe excepted. At Uranus, all names are derived from English literary (Shakespearean) sources, in keeping with the regular satellites, with no orbital subdivisions. At Neptune the small outer satellites are named after the 50 Nereids but generally follow the jovian scheme of “$a$” and “$o$” endings for prograde and retrograde satellites, with an “$e$” ending reserved for those with unusually high inclinations.

3. PHYSICAL PROPERTIES

The irregular satellites inhabit a unique niche in the solar system. They lie in an otherwise mostly empty region for stable small solar system bodies. They may be some of the only small bodies remaining that are still relatively near their formation locations within the giant planet region rather than being incorporated into the planets or ejected from the area. Their compositions may be intermediate between the rocky main-belt asteroids and volatile-rich TNOs.

3.1. Albedos and Size Distribution

Jupiter’s largest irregular satellites have very low albedos of about 0.04 to 0.05, which, along with their colors, are consistent with dark C-, P-, and D-type carbon-rich asteroids in the outer main belt (Cruikshank, 1977) and very similar to the jovian Trojans (Fernandez et al., 2003). The Cassini spacecraft obtained resolved images of Himalia that showed it to be an elongated object with axes of 150 × 120 km and an albedo of ~0.05 (Porco et al., 2003). Himalia’s relatively irregular shape makes it typical of small objects and satellites with internal pressures so low that they do not assume a hydrostatic shape [e.g., satellites with volumes $\leq 10^7$ km$^3$ exhibit axial ratios significantly greater than 1, according to Thomas et al. (1986)].

Saturn’s Phoebe has an average albedo of about 0.08 (Simonelli et al., 1999). Cassini obtained high-resolution images of Phoebe that showed it to be intensively cratered with many high albedo patches — possibly fresh ice exposures — located on crater walls (Porco et al., 2005). Further discussion of these results may be found in section 6.1. From Voyager data, Neptune’s Nereid was found to have an albedo of 0.16 (Thomas et al., 1991). These albedos for Phoebe and Nereid are more similar to the higher albedos found for large KBOs (Grundy et al., 2005; Cruikshank et al., 2005).

When measured to a given size, the populations and size distributions of the irregular satellites of each of the giant planets appear to be very similar (Fig. 2) (Sheppard et al., 2006). In order to model the size distributions we use a differential power law of the form $n(r)dr = \Gamma r^{-q}dr$, where $\Gamma$ and $q$ are constants, $r$ is the radius of the satellite, and $n(r)dr$ is the number of satellites with radii in the range $r$ to $r + dr$.

In the size range of $10 < r < 100$ km the irregular satellites of all four giant planets appear to have shallow power-
law size distributions with $q = 2$. Because of their proximity to Earth, only Jupiter and Saturn have known irregular satellites with $r < 5 \text{ km}$. These smaller satellites appear to follow a steeper power law ($q > 3.5$), which may be a sign of collisional evolution. In the Kuiper belt the size distribution is not well known at sizes similar to the irregular satellites, but KBOs with $r > 50 \text{ km}$ appear to have a much steeper power law ($q \sim 4$) than the irregular satellites. This distribution appears to become shallower for $r < 30 \text{ km}$, but more data are needed to determine a reliable Kuiper belt size distribution (Trujillo et al., 2001; Gladman et al., 2001b; Bernstein et al., 2004; Elliot et al., 2005; see also the chapter by Petit et al.).

If we do not include Triton, the largest irregular satellite at each planet is ~100 km in radius, with about 100 irregular satellites expected around each planet with radii larger than 1 km. This similarity between all four giant planet irregular satellite systems is unexpected considering the different formation scenarios envisioned for the gas giants vs. the ice giants (see section 5). Either the capture mechanism(s) for irregular satellites are independent of the planet formation process (Jewitt and Sheppard, 2005), or their size distributions are dominated by subsequent collisional evolution (see section 5.3).

### 3.2. Broadband Colors

Colors of the irregular satellites are neutral to moderately red (Tholen and Zellner, 1984; Luu, 1991; Rettig et al., 2001; Maris et al., 2001; Grav et al., 2003, 2004; Gray and Bauer, 2007). Most do not show the very red material found in the distant Kuiper belt (Fig. 3). The jovian irregular satellites’ colors are very similar to those of the carbonaceous C-, P-, and D-type outer main-belt asteroids (Degewij et al., 1980), as well as to the Trojans and dead comets. Colors of the Jupiter irregular satellite dynamical groupings are consistent with, but do not prove, the notion that each group originated from a single undifferentiated parent body. Optical colors of the eight bright outer satellites of Jupiter show that the prograde (or Himalia) group appears redder and more tightly clustered in color space than the retrograde irregulars (Rettig et al., 2001; Grav et al., 2003). Near-infrared colors recently obtained for the brighter satellites...
agree with this scenario and suggest that the jovian irregulars’ colors are consistent with D- and C-type asteroids (Sykes et al., 2000; Grav and Holman, 2004).

The saturnian irregular satellites are on average redder than Jupiter’s but still do not show the very red material observed in the Kuiper belt. Buratti et al., (2005) show that the color of the dark reddish side of Iapetus is consistent with dust from the small outer satellites of Saturn but not from Phoebe. Interestingly, most of Saturn’s smaller irregular satellites do not appear to have similar spectrophotometry to the neutral-colored Phoebe (Grav et al., 2003; Buratti et al., 2005), with the exception of Mundifari (Grav and Bauer, 2007). Although the Saturn irregulars do not show such obvious dynamical groupings in semimajor axis and inclination phase space as those at Jupiter, the optical colors and phase curves of several prograde saturnians do seem to be correlated for objects of similar inclination (Bauer et al., 2006; Grav and Bauer, 2007). The retrograde satellites, on the other hand, exhibit a wide range of optical colors with few strong orbital correlations (Grav and Bauer, 2007), consistent with several different parent bodies.

The irregular satellites of Uranus show a wide range of colors, especially in V–R (Maris et al., 2001; Romon et al., 2001; Grav et al., 2004), but little else can be said at present. There are only very limited observational data for Neptune’s irregulars, but to date they also do not appear to show the extreme red colors seen in the Kuiper belt.

Figure 3 shows that in general the irregular satellites, the Jupiter and Neptune Trojans, and cometary nuclei have similar colors that are quite different from the overall color distribution of Centaurs and KBOs. It has also been well established that the Kuiper belt appears to have two color populations, with the dynamically excited objects being less red on average (Tegler and Romanishin, 2000; Trujillo and Brown, 2002). Sheppard and Trujillo (2006) suggest that all the “dispersed populations,” i.e., those objects that are currently on stable orbits but were likely transported and trapped there, were derived from a similar location in the solar nebula because of their similar moderately red colors. These dispersed objects include the irregular satellites of the four giant planets as well as the Jupiter and Neptune Trojans and the dynamically “hot” — or excited — population of KBOs. If this scenario is true, it would suggest that all these populations were transported and trapped in their current locations at a similar time in the early solar system.

3.3. Spectra and Surface Compositions

Because the vast majority of the irregular satellites are very faint, little has been done in terms of spectroscopy except for the largest few objects. Optical and near-infrared spectra of the brighter jovian satellites are mostly featureless with moderately red slopes (Luu, 1991; Brown, 2000; Jarvis et al., 2000; Chamberlain and Brown, 2004; Geballe et al., 2002). Jarvis et al. (2000) find a possible 0.7-µm absorption feature in Jupiter’s Himalia and attribute this to oxidized iron in phyllosilicate minerals, which are typically produced by aqueous alteration. Cassini obtained a largely featureless near-infrared spectrum of Jupiter’s Himalia (Chamberlain and Brown, 2004). The neutral to moderately red linear spectra of Jupiter’s irregular satellites, like their albedos and colors, are consistent with those of dark, carbonaceous outer main-belt asteroids.

The one irregular satellite at Saturn and one at Neptune with measured spectra appear to be remarkably different from Jupiter’s irregular satellites. Saturn’s Phoebe and Neptune’s Nereid both show volatile-rich surfaces with strong water-ice signatures (Brown et al., 1998; Owen et al., 1999). In addition to almost ubiquitous water ice, the near-infrared spectra of Phoebe obtained by Cassini show ferrous-iron-bearing minerals (perhaps phyllosilicates), bound water, trapped CO₂, organics, and possibly nitriles and cyanide compounds (Clark et al., 2005). Phoebe’s diverse, volatile-rich surface suggests strongly that this object was formed beyond the main asteroid belt and that it may be more similar in composition to cometary nuclei. Further discussion is deferred to section 6.1.

4. DYNAMICS

4.1. Short-Term Dynamics

The largest perturbations on irregular satellite orbits invariably come from the Sun. Solar tides place the upper limit on the distance between the planet and the satellite. In the simplest approximation, the planet’s gravity will be stronger than solar tides within its so-called “Hill sphere,” the radius of which is given by

\[ R_H = \left( \frac{m}{3M} \right)^{1/3} r \]

(2)

where m and M are the planet’s and the Sun’s masses, and r is the mean distance between them. However, the region of long-term stability is somewhat smaller than the Hill sphere, and will be discussed in section 4.2.

On the other hand, most irregulars are sufficiently far from their parent planets not to be appreciably affected by the planet’s oblateness. The critical distance for the transition between oblateness-dominated and Sun-dominated dynamics is

\[ a_c = \left( \frac{2J_2 \frac{m}{M} R^2 r^3}{M} \right)^{1/5} \]

(3)

where J₂ and R are the planet’s second gravitational moment and radius. When we calculate this distance for Jupiter, Saturn, Uranus, and Neptune, we get 2.3, 2.5, 1.4, and 1.8 million km, respectively, although these distances are increased by ~30% when allowance is made for the contri-
bution to each planet’s effective $J_2$ by its equatorial, regular satellites. With the notable exception of Triton, all the satellites that were likely captured from heliocentric orbit have $a > a_c$.

The most important short-period solar perturbation of the orbit of a distant satellite is ejection. It is associated with the argument $2\lambda' - 2\sigma$, where $\lambda'$ is the solar true longitude and $\sigma$ is the satellite’s longitude of pericenter of center. The period of this perturbation is somewhat longer (or shorter, for retrograde satellites) than half the planet’s year. Ejection induces variability in all orbital elements, but its most notable effect is on the eccentricity. The eccentricity is largest when the Sun is along the line of the moon’s apsides, and lowest when the Sun is perpendicular to that line (this is again reversed for retrograde orbits). Lunar ejection is well described by Brouwer and Clemence (1961), and Čuk and Burns (2004b) have applied ejection to high-e and high-i orbits.

The Lidov-Kozai (LK) (Lidov, 1962; Kozai, 1962) mechanism is an extremely important consequence of solar perturbations, but was not discovered until the Space Age. This perturbation affects only objects with significantly inclined orbits. In the simplest treatment of the LK mechanism (e.g., Innanen et al., 1997), the interactions are purely secular, and the planet’s eccentricity has no direct role in the dynamics. The only relevant angle is the argument of pericenter $\omega$, which describes the orientation of the satellite’s line of apsides relative to its line of nodes on the planet’s orbit. During the precession cycle of $\omega$, eccentricity and inclination are coupled in such a way that the normal component of the angular momentum, $H = \sqrt{1 - e^2} \cos i$, is conserved. Eccentricity is at maximum (and inclination at minimum) when the line of apsides is perpendicular to the line of nodes ($\omega = 90^\circ$ or $\omega = 270^\circ$), while the minimum $e$ and maximum $i$ coincide with $\omega = 0^\circ$ or $\omega = 180^\circ$. Satellites with inclinations above 39.2° and below 140.8° can exhibit librations of $\omega$ about 90° or 270°. The first real satellites found to be in such a resonance were Saturn’s Ijiraq and Kiviuq (Vashkov’yak, 2001; Čuk et al., 2002; Nesvorný et al., 2003), but LK librators are now known around all four giant planets, including Euporie and Carpo at Jupiter, Margaret (S/2003 U3) at Uranus, and possibly Sao (S/2002 N2) and Neso (S/2002 N4) at Neptune (Carruba et al., 2004).

Finally, the traditional secular interactions involving the longitudes of pericenter ($\sigma$) of both the perturber and perturbee also can have a significant effect on the irregulars. Most distant irregulars (especially the retrograde jovians) show a perturbation in eccentricity that is governed by the angle between the satellite’s and the planet’s (or apparent solar) line of apsides, $\Psi = \sigma - \sigma'$, where $\sigma'$ is the planet’s longitude of perihelion. While such perturbations are generally weaker than the ejection and LK mechanisms, some irregular satellites are known to be in secular resonance, i.e., the argument $\Psi$ librates for some or all of the time. Jupiter’s Pasiphae (Whipple and Shelus, 1993) and Sinope (Saha and Tremaine, 1993) have $\Psi$ librating around 180°, while Saturn’s Sianaq and Narvi exhibit complex circulations of $\Psi$ (Čuk et al., 2002; Nesvorný et al., 2003; Čuk and Burns, 2004b). While Pasiphae’s librations can be very long-lived, none of these resonances appears to be primordial; instead, the critical argument undergoes chaotic variation over many millions of years (Nesvorný et al., 2003).

4.2. Long-Term Stability

Plate 8 clearly shows two trends among the irregular satellite orbits: They avoid high inclinations (60° < $i$ < 130°), and the prograde satellites are never found outside about a third of the Hill sphere, while the retrogrades extend out to about $R_H/2$.

The paucity of high-inclination satellites at Jupiter was explored in detail by Carruba et al. (2002). They show analytically that the LK mechanism will cause an initially high-inclination, circular orbit to become very eccentric at certain times during its secular cycle. This can lead to the orbit intersecting those of the Galilean satellites for a part of the irregular’s precession period, leading to an almost certain collision over the age of the solar system. Carruba et al. (2002) also performed numerical simulations for a grid of high-i orbits and showed that the range of inclinations unstable to the above effect is in reality even wider than predicted by analytical theory, due to the perturbations from the other three giant planets. Nesvorný et al. (2003) performed analogous integrations for the other three giant planets, with similar results.

Carruba et al. (2002) also find that the center of the “high-inclination hole” is slightly shifted away from 90° toward the retrograde region. This stems from the apsidal precession not being symmetric for prograde and retrograde bodies, despite LK theory predicting identical behavior. The asymmetric term is principally caused by the incomplete averaging of the ejection inequality (section 4.1). This extra term was first found to have major effects on the precession of the Moon by Clairaut in the eighteenth century [it almost doubles the predicted precession rate based strictly on secular perturbations (Baum and Sheehan, 1997)]. The “secular ejection” term always causes additional prograde precession of the apsides, which accelerates and decelerates the precession of prograde and retrograde orbits, respectively. Čuk and Burns (2004b) derived this term for orbits of any eccentricity and inclination using classical celestial mechanics. Beaugé et al. (2006) further improved the secular theory for irregulars using a more rigorous approach based on Hori (1966).

Averaged ejection also leads to the reduced stability of prograde orbits. The feedback between the solar perturbation and apsidal precession of prograde orbits leads to stronger coupling and large swings in eccentricity. At sufficiently large e, the apsidal precession rate will become comparable to the Sun’s mean motion, turning the ejection inequality into a destabilizing resonance (Nesvorný et al., 2003). This does not happen for retrograde orbits because their basic
Nevertheless, the nebula still needs to disappear on timescales if the capture has to happen during a single passage. In a much less dense gas environment than would be possible, a few tens of orbits, its permanent capture can be achieved. With the smallest body first going through a temporary capture lasting a time and leading to permanent capture of passing small bodies. However, this scenario requires the future satellite to share the Hill sphere of the planet with many Earth masses of accreting gas. It can be shown (Nesvorný et al., 2007) that the effects of aerodynamic drag on a small body will generally be more important than those of the planet’s changing mass, making the above scenario unrealistic.

Pollack et al. (1979) proposed that it was this aerodynamic drag in the collapsing envelope that produced the capture. However, since the orbital in-spiraling due to drag does not stop with capture, this method also requires a relatively rapid collapse of the circumplanetary nebula in order to remove the gas. Since gas drag affects smaller objects more strongly, surviving objects must have just the right size in order to be captured by drag but evolve slowly enough to outlive the collapse of the gaseous envelope. Additionally, any evolution by gas drag must have happened early, before the formation of dynamical families, as there is no indication of any relationship between size and either semi-major axis or eccentricity for the irregulars (e.g., smaller objects might be expected to be affected more strongly by gas drag), as illustrated in Fig. 4.

Čuk and Burns (2004a) tested the Pollack et al. (1979) scenario numerically and found that it could work for certain satellites, with the largest jovian irregular, Himalia, being the most likely candidate for such a capture. If the small body first goes through a temporary capture lasting a few tens of orbits, its permanent capture can be achieved in a much less dense gas environment than would be possible if the capture has to happen during a single passage. Nevertheless, the nebula still needs to disappear on timescales of $10^4$–$10^5$ yr in order for a Himalia-like body to survive, as there is no other way to stop its orbital evolution.

In principle, resonances can either capture a decaying orbit or lift its pericenter (by giving the body in question a negative eccentricity kick) and therefore save the irregulars from collapse. Unfortunately, the only resonance that is currently strong enough to induce large changes in the satellite’s eccentricity is the “great inequality,” in which the pericenters of prograde saturnians are caught in the 1:2 commensurability with the 900-yr near-resonant perturbation of Jupiter and Saturn. This perturbation arises from the proximity of these planets’ orbits to their mutual 2:5 mean-motion resonance. Čuk and Burns (2004b) find that this resonance can induce large secular variations in a satellite’s eccentricity that eventually lead to instability, and note that resonances of this type could have had important effects on irregular satellites during planetary migration (e.g., Hahn and Malhotra, 1999; cf. Carruba et al., 2004).

Tsiganis et al. (2005; see also chapter by Morbidelli et al.) proposed that Jupiter and Saturn crossed their mutual 1:2 resonance at some point in the past. Around the time of this resonance crossing, their irregular satellites would be subject to rapidly changing perturbations similar to the great inequality, but likely stronger. Čuk and Gladman (2006) postulated that continuous gas-drag capture and decay of bodies caught in the inner disk (where the regular satellites are now) would provide a steady-state population of bodies on continuously decaying eccentric orbits that can then
have their pericenters raised by the resonance passage. They numerically tested this hypothesis and found that a significant fraction of very high-e satellites of Saturn would have their eccentricities lowered by such a resonance passage. Their results show that the a and i distributions of bodies so stabilized strongly resemble the known irregular satellites, while the theoretical eccentricities are a little too high compared to the observed ones. A similar mechanism could also help capture permanent irregulars around Uranus and Neptune if they have ever crossed a resonance with Saturn (the orbits of the outer irregulars of Uranus also match the theoretical predictions for this mechanism). This mechanism does not work well for Jupiter, and it seems likely that other mechanisms helped capture its irregulars (see next section).

While Čuk and Gladman (2006) use the basic framework of Tsiganis et al. (2005), namely the Jupiter-Saturn resonance crossing, their model is not consistent with the full Nice model (see chapter by Morbidelli et al. and section 5.2 below). The scattering phase following the resonance crossing that Čuk and Gladman (2006) use to capture irregulars would also destroy any prior distant satellites, including those captured during the resonance crossing itself.

5.2. Purely Dynamical Capture Mechanisms

While a permanent capture is impossible in the gravitational three-body problem (Sun, planet, satellite), the presence of a fourth body can change this. Nesvorný et al. (2007) find that in the context of Tsiganis et al. (2005; also called the “Nice model,” see chapter by Morbidelli et al.), there are two distinct classes of encounters that could in principle result in irregular satellites. One involves encounters between a giant planet and a binary planetesimal, while the other involves encounters between two giant planets during the scattering phase of the Nice model, with the latter being more promising. One of the ice giants (Uranus or Neptune) typically has a few encounters with Saturn, while the two can have hundreds of mutual close fly-bys. While the Hill spheres of the two planets overlap, planetesimals passing through that region can get enough of a velocity

---

**Fig. 5.** A comparison between the orbits of objects captured in a numerical calculation (dots) with those of the known irregular satellites (solid triangles). (a,b) Satellites of Uranus; (c,d) satellites of Neptune. In total, 568 and 1368 stable satellites were captured at Uranus and Neptune, respectively. From Nesvorný et al. (2007).
kick relative to one of the planets to become its permanent satellites.

Nesvorný et al. (2007) tested this idea numerically, using the results of Gomes et al. (2005) to generate their initial conditions. They found this process to be efficient, producing a large number of stable irregulars at Saturn and an even larger number at Uranus and Neptune (see Fig. 5). The orbital distribution of these irregulars is mostly random within the stable region, except that very distant retrograde satellites rarely form. This is consistent with the present uranian irregular system, but only partially consistent with the saturnian one, where there are distinct “inclination groups” (Gladman et al., 2001a) (see section 2.3). While the overall mass of captured irregulars in their model is more than adequate to account for the present irregulars, the observed size distribution is much shallower (has fewer small bodies) than that of present-day TNOs. This might be a consequence of a different past size distribution of TNOs, observational bias, or other unmodeled processes.

Since Jupiter does not experience any close encounters with the other planets in the Nice model, none of its irregular satellites could have been captured by this mechanism. Since the mechanism of Čuk and Gladman (2006) cannot explain most jovians either, there is a discrepancy between theoretical expectations (which predict that jovians should be different) and observations of similar size distributions for irregular satellites around all four giant planets (cf. section 3.1). Either this similarity is just a coincidence or is due to significant collisional evolution, as discussed below, or the models of both Čuk and Gladman (2006) and Nesvorný et al. (2007) need modification, with the overall “Nice model” possibly requiring some fine-tuning.

5.3. Collisional Evolution and Families

The osculating orbits of irregular satellites are not constant on century timescales due to gravitational perturbations from the Sun and the other planets (see section 4.1). To determine which irregular satellites have similar orbits and may thus share a common origin, more constant orbital elements must first be defined. Several numerical and analytical methods have been developed for this purpose. The simplest method is to integrate numerically the satellites’ orbits over a suitably long timescale and determine the mean of their semimajor axes, eccentricities, and inclinations (hereafter denoted ⟨a⟩, ⟨e⟩, and ⟨i⟩). This is the method adopted by Nesvorný et al. (2003), as well as by R. A. Jacobson in generating the widely used mean orbital elements listed on the JPL HORIZONS website (http://ssd.jpl.nasa.gov/?horizons). Precise analytic definition of constant orbit elements is more difficult to achieve but is more flexible and can be easily repeated at low CPU cost when the orbit determinations improve (Beaugé and Nesvorný, 2007).

Figure 6 illustrates the distribution of numerically defined ⟨a⟩, ⟨e⟩, and ⟨i⟩ for the jovian retrograde satellites. As first noted by Sheppard and Jewitt (2003) and Nesvorný et al. (2003), two groups of tightly clustered orbits are apparent around Ananke and Carme. A third group around Pasiphae has also been proposed, although in this case the statistical significance of the concentration is low due to a small number of known orbits in the cluster. These satellite groups are reminiscent of the distribution of orbits in the main asteroid belt, where disruptive collisions between asteroids produced groups of fragments sharing similar orbits [the so-called asteroid families (Hirayama, 1918; Zappalá et al., 1994)].

The dispersion of orbits in the Ananke and Carme groups (hereafter families) corresponds to ejection speeds of 50 m/s, which nicely corresponds to values expected for the collisional breakup of putative 50-km-diameter parent satellites. Conversely, the prograde Himalia group at Jupiter and the so-called inclination groups at Saturn (e.g., the Phoebe group; see section 2.3) would indicate much larger ejection speeds that may be difficult to reconcile with what we know about large-scale disruptive collisions (Nesvorný et al., 2003; Grav and Bauer, 2007). Therefore, either the orbits
in these groups evolved significantly after their formation (Christou et al., 2005) or they were captured separately; Christou et al. argue for the collisional origin of the Himalia group. These results are in broad agreement with photometric observations, described in section 3.2, which show that objects in the Himalia, Ananke, and Carme families have similar colors, indicating similar physical properties (as expected for breakup of mineralogically homogeneous objects), while those in the Phoebe group at Saturn show substantial color diversity (Grav et al., 2003; Grav and Bauer, 2007). As discussed in section 2.3 above, no unambiguous orbital clusters have been found at Uranus and Neptune, perhaps due to the smaller numbers of currently known irregular moons at these planets.

Nesvorný et al. (2003), following Kessler (1981), calculated the rates of disruptive collisions between irregular moons. They found that (1) the large irregular moons must have collisionally eliminated many small irregular moons, thus shaping their population to the currently observed structures; (2) Phoebe’s surface must have been heavily cratered by impacts from an extinct population of saturnian irregular moons, much larger than the present one; and (3) disruptive collisions between jovian irregular moons cannot explain their orbital groupings. The first of these findings may account for the similarity in the size distributions noted in section 3.1.

The current impact rate on these moons from kilometer-sized comets and escaped Trojan asteroids is negligible (Nakamura, 1993; Zahnle et al., 2003). It has therefore been proposed that the origin of the Carme and Ananke families (and the Pasiphae family, if confirmed) dates back to early epochs of the solar system when impactors were more numerous (Sheppard and Jewitt, 2003). Nesvorný et al. (2004) analyzed the scenario whereby the satellite families form early by collisions between large parent moons and planetesimals. They found that the Ananke and Carme families at Jupiter could have been produced by these collisions unless the residual disk of planetesimals in heliocentric orbit was already severely depleted when the irregular satellites formed. Conversely, they found that formation of the Himalia group of prograde jovian satellites by the same mechanism was unlikely unless a massive residual planetesimal disk was still present when the progenitor moon of the Himalia group was captured. These results help to place constraints on the mass of the residual disk when satellites were captured, and when the Ananke and Carme families formed. Unfortunately, these constraints also depend sensitively on the assumed size-frequency distribution of planetesimals in the disk at 5–30 AU.

6. INDIVIDUAL OBJECTS

6.1. Phoebe

To date, Phoebe is the only irregular satellite, besides Triton, that has been studied in detail by spacecraft. Low-resolution Voyager images showed it to be an irregularly shaped body, roughly 100 km in radius, with some brightness variations on its surface. The Cassini flyby of December 12, 2004, provided the first close look at an object of this type. In addition to providing information on this satellite’s volatile-rich surface composition, the Cassini data on Phoebe’s shape and mass determined from tracking data provide a value for its mean density, a key constraint on its bulk composition.

Phoebe’s mean radius is 106.6 ± 1 km and its mean density is 1630 ± 45 kg m⁻³ (Porco et al., 2005). This density suggests a bulk composition consisting of a mixture of water ice and silicate, with the proportions depending on the amount of porosity present in this small body. Porco et al. note that even for zero porosity, Phoebe’s density is higher than that of the regular satellites (Triton’s uncompressed density is ~1500 kg m⁻³, while the average density of the smaller icy satellites is only ~1300 kg m⁻³), consistent with its assumed origin from outside the forming Saturn system. Its actual material density is probably even higher since it is plausible that Phoebe has significant bulk porosity, due to the low pressures in its interior (~4 MPa). Jupiter’s moon, Amalthea, for instance, is about the same size as Phoebe, but with a density of 857 ± 99 kg m⁻³ (Anderson et al., 2005).

Johnson and Lunine (2005) calculate that for a porosity of only 15%, Phoebe’s material density would be similar to the uncompressed densities of Pluto and Triton (~1900 kg m⁻³). This would be consistent with an origin in the outer protoplanetary nebula with about 70% of the carbon in the form of CO, given current revised solar abundances of C and O. They point out, however, that pure solar composition equilibrium condensation does not easily explain the low densities of the regular satellites. Another uncertainty is the amount of carbon in solid form during the condensation process. Given these uncertainties, perhaps the best general conclusion about Phoebe’s composition is that it is probably a body with at least modest bulk porosity and a material density indicating a silicate-rich composition compared with Saturn’s regular satellites. Combined with the observed water ice and volatiles on its surface (Clark et al., 2005), this suggests that it formed originally in the outer parts of the solar nebula from a reservoir of material similar to that which formed Pluto and Triton.

6.2. Triton and Nereid

Triton and Nereid satisfy some but not all criteria for being irregular. Triton’s orbit is retrograde, but close to Neptune and circular, while that of Nereid is large and eccentric, although Nereid might not be a captured body. McCord (1966) and McKinnon (1984) proposed that Triton is a captured satellite, whose originally eccentric orbit was circularized due to tidal dissipation within Triton. Goldreich et al. (1989) proposed that Triton was captured from heliocentric orbit by a collision with a preexisting satellite, and its initial high-eccentricity orbit then evolved due to tidal dissipation. McKinnon and Leith (1995) proposed that Triton was captured and evolved by gas drag, but this hypothesis has been less widely accepted than collisional capture. Agnor and Hamilton (2006) recently proposed a three-
body capture scenario for Triton. They suggest precapture Triton may have been a member of a binary whose disruption during a Neptune encounter led to Triton’s capture and its companion’s escape.

Neptune almost certainly had preexisting regular satellites, likely similar to those of Uranus (with a total mass of about 40% that of Triton). Čuk and Gladman (2005) show that the largest (hypothetical) satellites of Neptune would collide with each other after only \(10^3\) yr if perturbed by an eccentric and inclined early Triton, and then be ground into a disk. This disk would largely be accreted by Triton, causing rapid orbital decay, with final circularization due to tidal dissipation. In this picture, Triton is not simply a captured TNO, but an amalgamation of a large captured object and a significant amount of regular satellite material.

Goldreich et al. (1989) suggested that Nereid was a regular moon of Neptune, scattered onto an irregular orbit by the newly captured Triton. This agrees with Nereid being prograde, as well as with its icy spectrum (Brown et al., 1998), and opens the possibility that there could exist other, smaller bodies in similar orbits. Even after their orbits decoupled, Triton would have perturbed Nereid (and other relatively close-in irregulars) more strongly than the Sun. Čuk and Gladman (2005) find that test particles perturbed by a massive interior body with large e and i can oscillate between high-i prograde and retrograde orbits. Therefore, high-i retrograde objects could also be fragments of regular satellites.

Acknowledgments. The work described in this paper was carried out with support from NASA’s Planetary Astronomy program (P.D.N.) and the National Science Foundation (D.N.). S.S.S. was supported by NASA through Hubble Fellowship grant #HFF-01178.01-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. M.C. acknowledges support from the Canadian Institute for Theoretical Astrophysics (CITA) and the Natural Sciences and Engineering Research Council (NSERC) of Canada. A portion of this work was done at the Jet Propulsion Laboratory, California Institute of Technology, under a contract from NASA.

REFERENCES


search for Kuiper belt objects and Centaurs. II. Dynamical classification, the Kuiper belt plane, and the core population. *Astron. J.*, 129, 1117–1162.


Owen T., Cruikshank D., Dalle Ore C., Geballe T., Roush T., and


