Monthly Notices
of the
ROYAL ASTRONOMICAL SOCIETY

Mon. Not. R. Astron. Soc. **412,** 2137–2153 (2011) doi:10.1111/j.1365-2966.2010.18041.x

Collisional evolution of irregular satellite swarms: detectable dust around Solar system and extrasolar planets

G. M. Kennedy⋆ and M. C. Wyatt

Institute of Astronomy, University of Cambridge, Madingley Road, Cambridge CB3 0HA

Accepted 2010 November 16. Received 2010 November 15; in original form 2010 October 4

ABSTRACT

Since the 1980s it has been becoming increasingly clear that the Solar system's irregular satellites are collisionally evolved. The current populations are remnants of much more massive swarms that have been grinding away for billions of years. Here, we derive a general model for the collisional evolution of an irregular satellite swarm and apply it to the Solar system and extrasolar planets. The model uses a particle-in-a-box formalism and considers implications for the size distribution, which allows a connection between irregular satellite populations and predicted levels in the resulting dust cloud. Our model reproduces the Solar system's complement of observed irregulars well, and suggests that the competition between grain–grain collisions and Poynting–Robertson (PR) drag helps set the fate of the dust. In collision-dominated swarms most dust is lost to interplanetary space or impacts the host planet, while PR-dominated grains spiral in towards the planet through the domain of regular satellites. Because swarm collision rates decrease over time the main dust sink can change with time, and may help unravel the accretion history of synchronously rotating regular satellites that show brightness asymmetries, such as Callisto and Iapetus. Some level of dust must be present on au scales around the Solar system's giant planets if the irregular satellites are still grinding down, which we predict may be at detectable levels. We also use our model to predict whether dust produced by extrasolar circumplanetary swarms can be detected. Though designed with planets in mind, the coronagraphic instruments on *James Webb Space Telescope* (*JWST*) will have the ability to detect the dust generated by these swarms, which are most detectable around planets that orbit at many tens of au from the youngest stars. Because the collisional decay of swarms is relatively insensitive to planet mass, swarms can be much brighter than their host planets and allow discovery of Neptune-mass planets that would otherwise remain invisible. This dust could have been detected by *Hubble Space Telescope* Advanced Camera for Surveys (*HST* ACS) coronagraphic observations, and in one case dust may have already been detected. The observations of the planet Fomalhaut b can be explained as scattered light from dust produced by the collisional decay of an irregular satellite swarm around a $\sim$10 $M_\oplus$ planet. Such a swarm comprises about 5 lunar masses worth of irregular satellites. Finally, we briefly consider what happens if Fomalhaut b passes through Fomalhaut's main debris ring on a coplanar orbit, which allows the circumplanetary swarm to be replenished through collisions with ring planetesimals. This scenario, in which the planet is at least of the order of an Earth mass, may be ruled out by the narrow structure of the debris ring.

Key words: planets and satellites: detection – planets and satellites: general – circumstellar matter – planetary systems.

1 INTRODUCTION

With a penchant for retrograde, barely bound, high-eccentricity orbits and flatter-than-usual size distributions, the irregular satellites are one of the Solar system's rebel populations. Their presumed capture into these unusual orbits around the Solar system's giant planets has long been a puzzle. Because irregulars exist at all four outer planets, capture mechanisms specific to formation of gas giants like Jupiter and Saturn (e.g. Heppenheimer & Porco 1977; Pollack, Burns & Tauber 1979) are not general enough. Dynamical mechanisms, which do not rely on the existence or growth of large

⋆E-mail: gkennedy@ast.cam.ac.uk

gaseous atmospheres, can be applied to gas and ice giants alike and are therefore favoured (e.g. Colombo & Franklin 1971). Nesvorný, Vokrouhlický & Morbidelli (2007) recently proposed a different dynamical mechanism as part of a unified model of outer Solar system formation (Gomes et al. 2005; Tsiganis et al. 2005), where irregulars are captured during a period of instability via planet–planet interactions.

An apparent weakness of the planet–planet interaction mechanism is that the irregular satellites should share the same size distribution as Jupiter's Trojan asteroids, which in this model were captured at the same time from the same population (Morbidelli et al. 2005). In fact the differences in the Trojan and irregular satellite size distributions are marked, with irregular satellites being much flatter for sizes larger than about 10 km. To overcome this hurdle, Bottke et al. (2010) showed that while the size distribution was indeed *initially* steeper like the Trojans, 4.5 Gyr or so of collisional evolution is sufficient to reduce primordial irregular satellite populations to a size distribution that matches those currently observed. This result implies that there were previously more irregular satellites and perhaps most significantly that copious amounts of dust were produced during the depletion of these satellites.

The evidence for collisional evolution of irregular satellites has been mounting for some time. Kessler (1981) showed that the four prograde irregulars known to orbit Jupiter at the time had a relatively short collisional lifetime. The advent of large-format CCD surveys since the turn of the century has seen a dramatic increase in the number of irregulars and made further theoretical advances possible (e.g. Gladman et al. 2001; Sheppard & Jewitt 2003; Holman et al. 2004; Jewitt & Haghighipour 2007). Nesvorný et al. (2003) noted that irregular satellites around planets closer to the Sun have larger orbits (in Hill radii), suggesting that the lack of satellites closer to Jupiter is due to their erosion through collisions, which proceed at a faster rate closer to the planet. Using numerical integrations to derive average orbital elements, they also proved the existence of collisional families (see also Gladman et al. 2001). This latter discovery is particularly important, because the current collision rate amongst the irregular satellites is in some cases too low to explain their existence (Nesvorný et al. 2003). The inference is again that the number of irregular satellites, and thus their collision rate, was much greater in the past and that they decayed through collisions to the current level.

Coinciding with these theoretical advances was the discovery by direct imaging of Fomalhaut b (Kalas et al. 2008), an extrasolar planet predicted to exist based on the elliptical orbit of Fomalhaut's circumstellar debris ring (Kalas, Graham & Clampin 2005; Quillen 2006). The ring structure suggests that the planet is less than $3\,M_{\rm Jup}$, though the planet could be much less massive (Chiang et al. 2009). This discovery appears unrelated to irregular satellites, but the inability of Hubble, Keck and Gemini photometry to pin down whether the planet looks like a planetary atmosphere or reflected starlight provides the link.

While the planets discovered to orbit HR 8799 appear to be consistent with ∼1000 K substellar-mass objects (Marois et al. 2008), multiwavelength photometry of Fomalhaut b appears bluer than expected for a 200-Myr-old gas-giant planet. Specifically, Fomalhaut b has so far defied detection at wavelengths longer than 1 μm, leading Kalas et al. (2008) to suggest that the spectrum is actually starlight scattered from an optically thick circumplanetary disc of about 20 Jupiter radii. Though such a scenario is plausible, we argue that dust produced by a swarm of colliding irregular satellites is also a possibility.

Given that the Solar system's irregular satellite complement decayed to its current state through collisions and the exciting possibility that Fomalhaut b may harbour the first circumplanetary dust seen outside the Solar system, the time seems right to consider whether such clouds of irregular satellites could be visible around extrasolar planets. In the following sections, we derive a simple model for the evolution of a circumplanetary satellite swarm and the all-important extrasolar observable – the dust. We compare our model with the Solar system irregulars, and comment on the fate and observability of dust. We then apply the model to circumplanetary swarms around extrasolar planets. Finally, we explore what kind of satellite swarm could exist around Fomalhaut b, and what constraints this proposed swarm puts on the planet mass.

2 MODEL OF A CIRCUMPLANETARY SWARM

The irregular satellite swarms described in this paper have not knowingly been detected around other planets that orbit other stars. Therefore, like the pre-1995 d of planet formation theory, we must take cues from the Solar system. However, based on experience gained from the surprising diversity of extrasolar planets, we should not assume that our irregular satellite complement is typical, or that extrasolar analogues should follow all the same rules.

Thankfully, some of the most important irregular satellite properties are dynamical and would have been discovered even if the Solar system had no irregulars. The main dynamical curiosity is their inclinations, which are all within about 60° of the ecliptic (but include retrograde orbits). This evacuation of near-polar orbits is due to solar and planetary perturbations, which drive the eccentricities of highly inclined orbits to such large values that they either encounter regular satellites or leave the Hill sphere (Carruba et al. 2002; Nesvorný et al. 2003).

Another constraint comes from the stability of circumplanetary orbits. Although Shen & Tremaine (2008) find that satellites out to a few Hill radii could survive the age of the Solar system around Uranus and Neptune, all currently known irregulars have orbits with semimajor axes $a_{\rm pl}$ less than half the Hill radius $R_{\rm H}$ (e.g. Sheppard & Jewitt 2003; Holman et al. 2004; Sheppard, Jewitt & Kleyna 2005)

$$R_{\rm H} = a_{\rm pl}(M_{\rm pl}/3\,M_\star)^{1/3}, \quad (1)$$

where $M_{\rm pl}$ is the planet mass and $M_\star$ is the stellar mass. On the sky, the giant planets' Hill radii span several degrees. The exact stability limit has a small inclination dependence in that retrograde orbits are stable at larger distances than the widest stable prograde orbits (Nesvorný et al. 2003).

In this section we outline a model for the collisional evolution of irregular satellite swarms. Because we want to make predictions of the only possible extrasolar observable – dust – we keep our model simple. There are many uncertainties in extrapolating a swarm of irregular satellites to a cloud of dust, such as the strength of satellites, the size distribution slope, and the minimum grain size, which at this stage make the development of a more complex collisional model largely unnecessary.

The next four subsections contain many equations that describe properties of a circumplanetary swarm. Readers looking for actual numbers may like to refer ahead to Table 1, which shows estimates of some properties for the Solar system giant planets.

2.1 Collisional mass-loss

In a steady-state collisional cascade the mass within a given size range decreases as these objects are destroyed in collisions, but is

replaced at the same rate by fragments created by destruction of larger objects. Mass is lost at the bottom end of cascade, usually by radiation forces that remove grains smaller than some minimum size. The evolution of the size distribution is therefore dictated by the collisional decay of the largest objects.

The result of such a collisional cascade would, in an ideal situation (cascade infinite in extent, strength independent of size), have a steady-state size distribution with a well-defined slope of $n(D) = KD^{2-3q}$ where $n(D)\mathrm{d}D$ is the number of satellites between D and $D + \mathrm{d}D$ and $q = 11/6$ (Dohnanyi 1969).

In fact the true distribution of circumstellar collisional cascades like the asteroid belt is expected to have different slopes in different size ranges due to the way strength depends on size (O'Brien & Greenberg 2003). Strength is typically described by the parameter $Q^\star_\mathrm{D}$, which is the kinetic collision energy per target mass needed to shatter *and disperse* the target, such that the largest remnant is half the mass of the original target. Such a collision is commonly termed 'catastrophic'.

Small objects are held together by their own material strength, and grow weaker with increasing size due to the increased likelihood of the presence of a significant flaw. To quote Benz & Asphaug (1994), 'Subdivide this same rock into 100 equal pieces and 99 of them are now stronger than the original, owing to the simple fact that they do not contain the one weakest flaw.' Above some transition size ($D_\mathrm{t} \sim 0.1$ km) bodies gain strength from self-gravity. The energy needed for catastrophic disruption now increases with size. Though objects may be shattered, extra energy is required to ensure that the fragments have sufficient escape speeds and are no longer bound. Gravity also limits fracture propagation within the material, thus adding strength (e.g. Benz & Asphaug 1999). This behaviour is usually modelled using complex numerical codes, and parametrized by a power law for each of the strength and gravity regimes. In fact $Q^\star_\mathrm{D}$ varies by about a factor of 10 over the range of impact parameters and there are similar differences between strength laws derived by different studies (e.g. Benz & Asphaug 1999; Stewart & Leinhardt 2009). Thus $Q^\star_\mathrm{D}$ and the resulting size distribution are the most uncertain inputs for our model.

We set planetesimal strength with the Benz & Asphaug (1999) law for ice at 3 km s^{-1}. For objects larger than $D_\mathrm{t} = 0.1$ (in km),

$$Q^\star_\mathrm{D} = 0.1\rho D^{1.26}/f_\mathrm{Q} \tag{2}$$

in J kg^{-1} where the mass density ρ is in kg m^{-3}. The strength dependence of small objects ($\propto D^{-0.39}$) is only used in setting the size distribution of objects smaller than D_t. Following Bottke et al. (2010), we allow objects to be weaker than the Benz & Asphaug (1999) law by including the factor f_Q (see also Levison et al. 2009). The strength law is similar to the Stewart & Leinhardt (2009) strength law when $f_\mathrm{Q} = 8$. Bottke et al. (2010) reproduce the Solar system's irregular satellite populations best when $f_\mathrm{Q} > 3$; so we set $f_\mathrm{Q} = 5$.

The size distribution of objects with such strength properties is expected to have a slope with $q_\mathrm{s} = 1.9$ at $D < D_\mathrm{t}$ (in the strength-dominated regime) and $q_\mathrm{g} = 1.7$ for $D > D_\mathrm{t}$ (in the gravity-dominated regime; O'Brien & Greenberg 2003). Although several wiggles are also expected in the distribution (Campo Bagatin et al. 1994; Durda, Greenberg & Jedicke 1998), here the size distribution is assumed to be continuous with the appropriate slopes (q_s or q_g) between the smallest objects of size D_min (in μm) and the largest objects participating in the collisional cascade of size D_c (in km). As we show later, D_c may be smaller than the largest object, which has size D_max.

With this two-phase size distribution, the conversion between the size distribution's surface area (σ_tot in au^2) and mass (M_tot in $\mathrm{M}_\oplus$) is

$$M_\mathrm{tot} = 0.0025\rho\sigma_\mathrm{tot}\,\frac{3q_\mathrm{s}-5}{6-3q_\mathrm{g}}\,D_\mathrm{c}^{6-3q_\mathrm{g}}\,D_\mathrm{t}^{3q_\mathrm{g}-3q_\mathrm{s}}\,D_\mathrm{min}^{3q_\mathrm{s}-5}, \tag{3}$$

where we assume $D_\mathrm{min} \ll D_\mathrm{t} \ll D_\mathrm{c}$. For $q_\mathrm{s} = 1.9$, $q_\mathrm{g} = 1.7$ and $D_\mathrm{t} = 0.1$ (in km), this relation simplifies to

$$M_\mathrm{tot} = 3.9 \times 10^{-6}\,\rho\sigma_\mathrm{tot}\,D_\mathrm{c}^{0.9}\,D_\mathrm{min}^{0.7}. \tag{4}$$

This mass only includes objects between D_min and D_c (not D_max) because we use M_tot below to calculate collision rates.

The collisional lifetime of satellites of size D_c can be calculated from the total mass using the particle-in-a-box approach. Here we follow Wyatt et al. (2010), who also take into account that objects of size D_c can be destroyed in catastrophic collisions by those down to a size $X_\mathrm{c}D_\mathrm{c}$ (assumed to be $\gg$0.1 km) where $X_\mathrm{c} = (2Q^\star_\mathrm{D}/v^2_\mathrm{rel})^{1/3}$ and the collision velocity v_rel is in m s^{-1}. The rate of catastrophic collisions is

$$R_\mathrm{cc} = 8.4 \times 10^{-5}\,\frac{6-3q_\mathrm{g}}{3q_\mathrm{g}-5}\,\frac{v_\mathrm{rel}C_1X_\mathrm{c}^{C_2}M_\mathrm{tot}}{\rho D_\mathrm{c}V} \tag{5}$$

in yr^{-1}, where V is the volume occupied by the satellites in au^3. The inverse of the collision rate is called the collision time t_c. The assumption of $D_\mathrm{t} \ll X_\mathrm{c}D_\mathrm{c}$ means that only q_g is needed for the collision rate. To obtain equation (5), it was necessary to integrate over the size distribution from $X_\mathrm{c}D_\mathrm{c}$ to D_c, which yields a function called $G(q, X_\mathrm{c})$ by Wyatt et al. (2007). In Wyatt et al. (2010) the approximation $G(11/6, X_c) \approx C_1X_c^{C_2}$ with $C_1 = 0.2$ and $C_2 = -2.5$ is used (the limit of small X_c). Substituting these values and $q_\mathrm{g} = 11/6$ yields equations (9) and (10) of Wyatt et al. (2010). Here, we find a numerical approximation for the function $G(q, X_c)$, which is within 10 per cent for the more physically plausible range $0.01 < X_c < 0.75$ and $1.7 < q < 2$, yielding $C_1 = 2.62(q - 1.66)$ and $C_2 = 2.70(0.98 - q)$. For $q_\mathrm{g} = 1.7$, $C_1 = 0.1$ and $C_2 = -1.9$.

Irregular satellites are assumed to orbit the planet at semimajor axes relative to the Hill radius in the range $\eta \pm \mathrm{d}\eta/2$ (we use $\mathrm{d}\eta = \eta/2$). The volume the satellites occupy is $V = 4\pi\eta_\mathrm{s}^2\mathrm{d}\eta_\mathrm{s}r_\mathrm{H}^3 \times 0.866$, where the extra factor of 0.866 accounts for the lack of near polar orbits.

The mean relative velocity of collisions is expected to be some fraction f_{v_rel} of the Keplerian velocity at η, which (in m s^{-1}) is

$$v_\mathrm{k} = 516\,M_\mathrm{pl}^{1/3}\,M_\star^{1/6}/(\eta\,a_\mathrm{pl})^{1/2}. \tag{6}$$

That fraction will depend on the eccentricities and inclinations of the satellite swarm. A simple estimate of the mean collision velocity comes from assuming circular and isotropic orbits, yielding $f_{v_\mathrm{rel}} = 4/\pi$ and typical impact velocities of ∼0.5–3 km s^{-1} for Solar system giant planets. These velocities are high enough that the impactor/target mass ratio for catastrophic collisions is small, so the energy lost in a collision is also small. Thus very little kinetic energy is lost in a typical collision and collisional damping is unimportant. Using the Monte Carlo eccentric ring model of Wyatt et al. (2010), we find that the mean collision velocity is similar for a realistic orbital distribution, and somewhat lower when eccentricities are introduced. The values vary between about 0.9 and 1.3, so we adopt $f_{v_\mathrm{rel}} = 4/\pi$ throughout. In fact orbits have a range of inclinations and eccentricities, and each collision has a different probability, which itself is a function of the relative velocity (Bottke et al. 1994).

Substituting our approximation for C_1 and C_2 for $q_\mathrm{g} = 1.7$ yields the rate of catastrophic collisions

$$R_\mathrm{cc} = 1.3 \times 10^7\,\frac{M_\mathrm{tot}M_\star^{1.38}f_{v_\mathrm{rel}}^{2.27}}{{Q^\star_\mathrm{D}}^{0.63}\rho D_\mathrm{c}M_\mathrm{pl}^{0.24}(a_\mathrm{pl}\,\eta)^{4.13}} \tag{7}$$

in yr^{-1}. As one expects, the rate is independent of $a_{\rm pl}$ for the same physical swarm (i.e. doubling $a_{\rm pl}$ halves η). The factors that largely set the collision rate are η, $D_{\rm c}$, $a_{\rm pl}$ and $f_{v_{\rm rel}}$. The strongest contributions are from η and $a_{\rm pl}$, which set the cloud volume and space density. The rate also depends strongly on the mean collision velocity because this speed sets both the rate at which an object sweeps through space and the number of impactors that result in a catastrophic disruption. Greater collision velocities mean that smaller impactors can destroy a given object, and smaller impactors are more numerous. Because $Q_{\rm D}^{\star}$ also depends on $D_{\rm c}$ (equation 2), the collision rate also depends strongly on $D_{\rm c}$. For the same total mass, larger $D_{\rm c}$ means fewer large objects, which are also stronger.

Curiously, the least important parameter is the mass of the planet itself. This result arises because in equation (7), $v_{\rm rel}$ is slightly less than linearly dependent on $M_{\rm pl}$, which nearly cancels with the linear dependence of volume on $M_{\rm pl}$. In simple terms, for fixed η and $M_{\rm tot}$ the increase in volume with planet mass works out to be slightly stronger than the increase in velocity, so the collision rate decreases slowly as the planet mass increases.

We validate our model with the eccentric ring model (Wyatt et al. 2010). Our collision rate is within a factor of 3 for a range of eccentricities with the best agreement for high values, sufficient for our purposes here considering that much larger uncertainty lies with assumed material properties and the resulting size distribution.

2.2 Radiation forces on dust

To derive the surface area in small grains (dominated by small objects) from the total mass (dominated by large objects) we need to know the size of the smallest grains that can survive in circumplanetary orbits. Two notable detections are micrometre-sized grains found orbiting at large (50–350 $R_{\rm Jup}$) distances from Jupiter (Krivov et al. 2002; Krüger et al. 2010), and the large tenuous ring found orbiting Saturn (Verbiscer, Skrutskie & Hamilton 2009). Both studies attribute material released by impacts from interplanetary grains as a likely source, though Verbiscer et al. note that debris from irregular satellite collisions impacting Phoebe could also be the cause.

As with grains orbiting a star, the effect of radiation forces on dust characterized by $\beta = F_{\rm radiation}/F_{\rm gravity}$ (both due to the star) plays the most important role in setting the minimum size of grains that survive on circumplanetary orbits (Burns, Lamy & Soter 1979). Other effects related to interaction with planetary magnetospheres (e.g. Horanyi 1996) play some part but are less important for the $\gtrsim$μm grain sizes and wide orbits considered here.

2.2.1 Radiation pressure

Radiation pressure is the radial component of the force, which in contrast to circumstellar orbits causes the orbits of dust grains to evolve. While semimajor axes remain constant, eccentricities oscillate with a period equal to the planet's orbital period, with a maximum that depends on β and the grain orbit. The maximum β of grains that survive in orbit around the planet with $e < 1$ have (Burns et al. 1979)

$$\beta_{\rm c} = v/3\,{\rm v}_{\odot} = 5.8 \times 10^{-3} M_{\rm pl}^{1/3}/(M_{\star}^{1/3}\eta^{1/2}), \quad (8)$$

where v is the velocity of a grain as it orbits the planet and ${\rm v}_{\odot}$ is the velocity of the planet as it orbits the star. For typical planets and irregular satellite orbits $\beta_{\rm c}$ is much smaller than the blowout limit of 0.5 for stellocentric orbits.

Because β for normal grains peaks where the star radiates most of its radiation, it might be possible for $\beta_{\rm c}$ to allow both large and very small grains to survive, with only grains in the peak being excluded. However, for submicrometre 'astronomical silicate' grains, those on the small side of the peak, the smallest grains have $\beta \sim 0.11$ (e.g. Gustafson 1994) and is higher for more massive stars, so grains smaller than the wavelength of typical stellar radiation will usually be ejected.

To convert $\beta_{\rm c}$ into a minimum size, we use $D_{\rm min} = (1150/\rho\beta_{\rm c})L_{\star}/M_{\star}$ in μm (e.g. Wyatt 2008a). The minimum size is therefore

$$D_{\rm min} = 2 \times 10^5 \frac{\eta^{1/2}\, L_{\star}}{\rho M_{\rm pl}^{1/3} M_{\star}^{2/3}} \quad (9)$$

in μm. For typical parameters, $D_{\rm min}$ is at least μm size. For Jupiter and Neptune, equation (9) yields 12 and 23 μm, an order of magnitude larger than the minimum (blowout) size for the same grains on circumsolar orbits.

2.2.2 Poynting–Robertson drag

An alternative to grain removal by radiation pressure is orbital decay due to Poynting–Robertson (PR) drag. The decay time-scale is similar to the heliocentric case

$$t_{\rm PR} = 530 a_{\rm pl}^2/(\beta M_{\star}) \quad (10)$$

in years (Burns et al. 1979). However, as noted above β for the smallest grains is typically much smaller than 0.5 and $t_{\rm PR}$ correspondingly longer.

For grains to spiral into the planet by PR drag they must avoid colliding with other grains first, which breaks them into smaller particles that are instead removed by radiation pressure. The competition between PR drag and collisions can be characterized by $\chi_{\rm PR} = t_{\rm PR}/t_{\rm col}$. Wyatt et al. (1999) showed that for the smallest grains this collision rate is roughly $t_{\rm per} r_{\rm dust}^2/(4\sigma_{\rm tot})$ for a flat disc with radial extent $r_{\rm dust} \pm r_{\rm dust}/4$ (where $t_{\rm per}$ is orbital period). Adapting this expression to an isotropic case results in a small change due to the greater cloud volume and faster collision velocities: $t_{\rm col,dust} = t_{\rm per} r_{\rm dust}^2/(4 f_{v_{\rm rel}} \sigma_{\rm tot})$, or using our parameters

$$t_{\rm col,dust} = 10^{-5} \frac{(\eta a_{\rm pl})^{7/2} M_{\rm pl}^{2/3}}{M_{\star}^{7/6}\sigma_{\rm tot}} \quad (11)$$

in years. Expressed in terms of our basic parameters, the ratio is

$$\chi_{\rm PR} = 4 \times 10^4 \frac{\rho D_{\rm min}\sigma_{\rm tot} M_{\star}^{7/6}}{a_{\rm pl}^{3/2}\eta^{7/2} M_{\rm pl}^{2/3} L_{\star}} \quad (12)$$

When this ratio is larger than unity, grains suffer collisions before their orbits have time to decay due to PR drag, and are subsequently removed by radiation pressure (termed 'collision dominated'). When this ratio drops below unity, grains spiral in towards the planet before they collide and may encounter any existing regular satellites as they do so ('PR dominated'). Substituting equation (9) for $D_{\rm min}$ yields

$$\chi_{\rm PR} = 8 \times 10^9 \frac{\sigma_{\rm tot} M_{\star}^{1/2}}{a_{\rm pl}^{3/2}\eta^3 M_{\rm pl}}. \quad (13)$$

With only the Solar system example to go by (see Section 3.1.1), it is impossible to tell whether the 'typical' satellite dust cloud will be collision or PR dominated. However, due to the amount of dust needed for an extrasolar irregular satellite swarm to be detectable ($\sigma_{\rm tot} \gtrsim 10^{-4}\,{\rm au}^2$), any observed extrasolar swarms will likely be collision dominated.

2.2.3 Summary

In order to estimate the minimum grain size we have necessarily made a number of simplifications. We have used expressions for radiation forces assuming low inclination orbits akin to Saturn's ring. In fact β_c varies with the orientation of the orbit relative to the solar direction and is the smallest for those with pericentre initially aligned with the solar direction. Equation (8) assumes a coplanar orbit with pericentre perpendicular to the solar direction and underestimates β_c for certain loss by a factor of 2–3 (Burns et al. 1979). We therefore overestimate the minimum grain size, and underestimate σ_{tot}.

Our prescription for D_{min} does not preclude detection of grains smaller than D_{min}. Grains below the minimum size may complete many orbits before their eccentricity exceeds unity. Also, the minimum grain size decreases with η so small bound eccentric grains with smaller planetocentric semimajor axes can be found anywhere within 2η. On a detailed level, D_{min} takes on a range of values and for the smallest grains with $\beta \sim \beta_c$, the orbit averaging used in deriving equation (8) breaks down. Using numerical simulations, Krivov et al. (2002) find that D_{min} for Jupiter is $\sim$1 μm (compared to our value of 12.4 μm, see Table 1). Considering that our value is overestimated by the factor of 2–3 noted above, the minimum size is probably 1 to a few μm, with differences in the assumed grain properties contributing some uncertainty.

While our simplifications are reasonable, they gloss over important aspects of grain dynamics. For example, Krivov et al. (2002) show circumplanetary dust clouds to have both size-dependent and prograde/retrograde orbit sensitive structure. Because non-gravitational forces cause small objects to deviate from the orbits of their parent bodies, such forces lead to effects beyond the scope of our model. Therefore, future work will need to consider how grain orbital evolution affects both the physical appearance of the cloud, and the underlying size distribution. For example, in a typical collision-dominated circumstellar disc the minimum (blowout) grain size is a single number, independent of radial distance. This fairly sharp truncation creates a wave in the size distribution due to the alternating lack and then overabundance of projectiles that destroy larger targets (Campo Bagatin et al. 1994). However, in a circumplanetary cloud the minimum grain size varies with circumplanetary distance (equation 9), which could lead to qualitatively different and spatially varying size distributions. In Section 3.1.2 we suggest that Jupiter's swarm may be PR dominated, which can lead to further differences at the small end of the size distribution. In addition, other non-gravitational forces known to be important for circumstellar dynamics (e.g. Yarkovsky effect; Burns et al. 1979) may be important for the evolution of circumplanetary orbits and consequent collisional evolution.

2.3 Time evolution

We now turn to the time evolution of disc properties. Assuming that the size distribution remains fixed and that mass is lost by catastrophic disruption of the largest objects, the disc mass remaining as a function of time is found by solving (e.g. Wyatt et al. 2007)

$$\frac{\mathrm{d}M_{tot}}{\mathrm{d}t} = -M_{tot} R_{cc}, \tag{14}$$

which yields

$$M_{tot}(t) = \frac{M_{tot}(0)}{1 + R_{cc}(0)\, t}, \tag{15}$$

where we use t or 0 to indicate variables at a particular time where necessary. The initial cloud mass is $M_{tot}(0)$ and $M_{tot}(t) = M_{tot}(0)/R_{cc}(0)\, t$ when $R_{cc}(0)\, t \gg 1$. Equation (4) shows that mass and surface area are linearly proportional for fixed D_c and D_{min} so that the surface area of small grains decays in the same way. Because the collision rate depends on the remaining mass, after a few collision times the remaining mass is independent of the initial mass (Wyatt et al. 2007).

2.3.1 Where are the most massive swarms?

There is an interesting interplay between the initial cloud mass and planet semimajor axis. For arbitrarily large initial masses, swarms at any distance and time are on the $1/t$ part of their evolution where the remaining mass is independent of the initial mass (called a 'collision-limited' disc by Heng & Tremaine 2010). Because these systems have a maximum remaining mass that depends on the collision time ($1/R_{cc}$) without the mass term, the remaining mass increases strongly with planet semimajor axis. In fact, $t_c \propto a_{pl}^{4.13}$ (i.e. more strongly than a_{pl}^2), so collision-limited swarms are brighter in scattered light for larger planet semimajor axes (for fixed planet mass, age, etc.).

Of course, the initial swarm mass cannot be arbitrarily large, so swarms around sufficiently distant planets will take some time to start colliding. Thus, for fixed initial mass, swarms around close planets will rapidly decay due to the short collision times, while more distant planets all have the same mass in satellites because the largest irregulars have not yet, or only just begun, to collide. Of these more distant planets, the one whose swarm has just started to suffer collisions is the one that receives the most stellar insolation and is brightest in scattered light.

The semimajor axis of the planet hosting this swarm can easily be worked out from the collision rate in equation (7), because a swarm that has just started to collide has $t = 1/R_{cc}$, therefore

$$a_{opt} = 50 \frac{M_\star^{0.33} f_{v_{rel}}^{0.55}}{M_{pl}^{0.06} Q_D^{\star\,0.15} \eta} \left(\frac{t M_{tot}(0)}{\rho D_c} \right)^{0.24} \tag{16}$$

in au. The 'opt' subscript indicates that this planet lies at the optimum distance to be detected at 'optical' wavelengths. This prescription for the brightest swarm is complicated by the fact that faint objects are harder to detect close to host stars, an issue we return to in Section 3.2. The same concept of an optimal distance applies to thermal emission, but is somewhat different because the cloud temperature changes with planet semimajor axis.

2.3.2 Stranding the largest objects

If collisional evolution proceeds for long enough, the mass in the swarm will drop to the point where it is comparable with the mass contained in a single largest object. Around this time the evolution of the largest objects changes from being reasonably well described by our particle-in-a-box formalism, to a regime where individual collisions and cratering are important (Bottke et al. 2010). In this regime, the largest objects are less likely to be destroyed due to their small number and the decreased number of potential destructors. While our model cannot take cratering or stochastic collisions into account, we can approximate the evolution by assuming that when the number of largest objects drops too low, they lose their connection with the rest of the size distribution and become 'stranded'. The size of the largest non-stranded object is of size D_c by definition, which decreases over time and leaves a relatively flat size

distribution of stranded objects between sizes D_c and D_{max}, whose evolution is halted due to a lack of would-be destructors.

We implement this simple approximation by assuming that an object of size D is stranded when the number of objects between sizes $D/2$ and D,

$$n(D/2 \rightarrow D, t) = \frac{K(2^{3q_g-3} - 1)(10^3 D)^{3-3q_g}}{3q_g - 3}, \quad (17)$$

drops to some number N_{str}. This number sets the normalization for the size distribution of stranded objects. This assumption sets the size distribution slope of stranded objects $q_{str} = 1$ because $n(D/2 \rightarrow D, t)$ is independent of D for $D_c > D > D_{max}$.

The first object is stranded at t_{nleft}, which can be calculated from the initial number of objects in this size range:

$$t_{nleft} = \frac{n(D_c/2 \rightarrow D_c, t = 0)}{R_{cc}(0)\, N_{str}}. \quad (18)$$

The difference between t_{col} and t_{nleft} is simply a measure of the initial number n of large objects, which take n/N_{str} collision times to become stranded. Like the remaining mass once collisions occur, this time is independent of the initial cloud mass.

At this point, D_c and D_{max} become distinct sizes, with D_{max} remaining fixed and D_c decreasing with time as smaller objects are stranded. The remaining planetesimal population decays at the collision rate for D_c size and strength objects and the mass remaining in the size distribution below D_c. Therefore, by substituting D_c for M_{tot} in equation (14), the evolution after t_{nleft} obeys

$$\frac{dD_c^3}{dt} = -D_c^3 R_{cc}, \quad (19)$$

where we have used $M_{tot} \propto D_c^3$ [for fixed $n(D_c/2 \rightarrow D_c)$]. The collision rate is for size D_c objects, on which both M_{tot} and $Q_D^\star$ depend, and varies as $R_{cc} \propto D_c^{1.2}$ (equations 2 and 7). Integrating equation (19) yields

$$D_c = \frac{D_{max}}{[1 + 0.4(t - t_{nleft})/t_{nleft}]^\alpha}, \quad (20)$$

where $\alpha = 1/1.2$. This evolution is illustrated in Fig. 1, which shows the time evolution of the top section of the size distribution and σ_{tot}. The size distribution initially decays straight down (i.e. K decreases) with D_c fixed. When there are only N_{str} largest objects left, D_c begins to decay as dictated by equation (20). The size distribution then moves to the left (smaller D_c), with both σ_{tot} and M_{tot} continuing to decay. The right-hand panel shows that σ_{tot} drops more quickly after the first objects are stranded, tending to $\sigma_{tot} \propto t^{-1.75}$ (using equations 4 and 20). Such evolution is potentially interesting, as a dust cloud is accelerated towards being PR dominated after stranding due to the stronger decrease in small grains with time. Entering the PR-dominated regime, the size distribution is effectively truncated at larger sizes, leading to an even faster decay of σ_{tot} than that shown in Fig. 1 (Dominik & Decin 2003).

This evolution is necessarily very simple because objects are only destroyed by catastrophic collisions in our model. However, at these late stages the mass released into the cascade by cratering may be as or more important (Bottke et al. 2010). Therefore, though our model is physically plausible, the actual evolution will depend on details such as the relative importance of catastrophic disruptions versus cratering or differences between prograde and retrograde populations. We treat N_{str} and α as free parameters when comparing our model with the Solar system irregulars in Section 3.1.

2.4 Observable quantities

Given the surface area in dust, we derive the flux density F due to the cloud and planet from both scattered light and thermal emission. Because distance and surface area can be in different units (e.g. m, pc, au), quantities in these equations have dimensions. We take Solar system planetary effective temperatures and radii from Cox (2000). The stellar flux at the planet is

$$F_\star = L_\star B_\nu(\lambda, T_\star) / \left(4\sigma_K T_\star^4 a_{pl}^2\right), \quad (21)$$

where σ_K is the Stefan–Boltzmann constant. The scattered light seen from Earth is (e.g. Collier Cameron et al. 2002)

$$F_{scat} = F_\star A g Q / \left(\pi d_{pl}^2\right), \quad (22)$$

where A is either the projected area of the planet or σ_{tot} for dust. The geometric albedo Q is assumed to be 0.08 for dust, similar to both the values for Kuiper belt objects, Jovian trojans and irregular satellites in the Solar system (Mueller et al. 2008; Stansberry et al. 2008; Fernández, Jewitt & Ziffer 2009) and to that inferred for the Fomalhaut dust ring (Kalas et al. 2005). For planets we use an albedo of 0.5. The phase function g is set to unity for the Solar system because planets are on exterior orbits. For extrasolar planets and swarms we set $g = 0.32$, the value for a Lambert sphere at maximum extension from the host star (e.g. Collier Cameron et al.

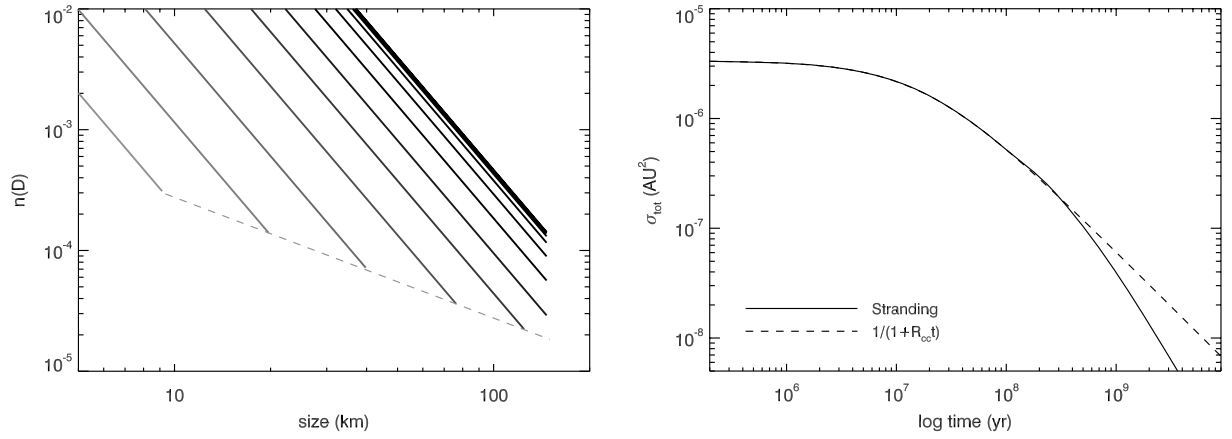

Figure 1. Example evolution of the size distribution (left-hand panel) and total surface area σ_{tot} (right-hand panel) for 10 Gyr. In the left-hand panel the lines become lighter for later times and are logarithmically spaced in time. The dashed line shows stranded objects. In the right-hand panel the evolution with stranding is shown as a solid line and the 'normal' $1/t$ evolution with no stranding as a dashed line.

2002). We use a blackbody estimate for the dust temperature $T_{\rm dust} = 278.3\,L_\star^{1/4}/a_{\rm pl}^{1/2}$ K. The distance $d_{\rm pl}$ is the distance to the system in the extrasolar case, or $a_{\rm pl}$ in the Solar system (i.e. roughly the distance for a Sun–Earth–planet angle of 90° for outer planets). For thermal emission,

$$F_{\rm th} = B_\nu(\lambda, T)\, A/d_{\rm pl}^2. \quad (23)$$

For dust we include a 'grey body' decrease in emission of $210/\lambda$ beyond 210 μm to account for inefficient emission from small grains (e.g. Wyatt 2008b). Though the actual spectrum depends on the grain properties and size distribution, this addition provides more realistic (sub)millimetre flux densities than a plain blackbody.

3 APPLICATIONS

In this section we apply our model to three different irregular satellite populations. We first compare our model evolution with the Solar system's complement of irregulars and make predictions of the current levels of dust. We suggest that the fate of dust changes over time, with implications for regular satellites such as Callisto and Iapetus. We then apply our model to possible satellite clouds around extrasolar planets. Finally, we apply our model to Fomalhaut b.

3.1 Solar system irregular satellites

In this section we compare our model with the Solar system's irregular satellites. We model a swarm with initial mass $M_{\rm tot}(0) = 0.01 M_{\rm moon}$. This mass ensures all swarms are collision limited and therefore does not influence the model. We set $D_{\rm c} = 150$ km for Jupiter and Uranus and $D_{\rm c} = 250$ km for Saturn and Neptune. We set the density similar to known irregular satellites $\rho = 1500\,{\rm kg\,m^{-3}}$ (e.g. Jewitt & Haghighipour 2007). Though we assume $f_{v_{\rm rel}} = 4/\pi$ at each planet, this number varies somewhat depending on the specific orbital properties of each swarm. We use the average η of known irregulars at each planet.

Fig. 2 shows a comparison of our model with the Solar system's irregular satellites.[1] We have combined the prograde/retrograde satellites for this distribution. The simplicity of our model means that we cannot account for the different evolution of prograde/retrograde populations, which appears to be important at Uranus (Bottke et al. 2010). Because all satellites at each planet are at a similar distance from Earth, these distributions are near-complete to the smallest observed size and need no correction (e.g. Sheppard et al. 2005).

We vary the normalization $N_{\rm str}$ and rate of decay after stranding power-law index α to obtain a by-eye fit. Variation of these parameters over a fairly small range allows an excellent match for Jupiter, Saturn and Uranus. Given the differences in the irregular populations at each planet and the simplicity of the prescription for stranding, some variation is expected. The first few columns of Table 1 show the η, $t_{\rm col}$ and $t_{\rm nleft}$ for each planet. The slower collisional evolution for more distant planets means that $t_{\rm nleft}$ is longer and that the size distribution at the current epoch turns up (i.e. stranded) at larger sizes. The minimum known satellite size increases for more distant planets, making the model comparison less certain for Uranus and Neptune. We do not know if their size distributions are similar to Jupiter and relatively flat to ∼8 km, or instead turn up at larger sizes like our model. If evolution at Uranus is similar to Jupiter and Saturn, our model predicts that the size distribution should turn up at a few tens of km due to the slower collisional evolution.

There are few Neptune irregulars to compare with, perhaps because many were depleted by Nereid (Nesvorný et al. 2003) and/or Triton (Goldreich et al. 1989; Ćuk & Gladman 2005). Nereid's low orbital inclination (7° relative to the ecliptic) and relatively close prograde orbit ($\eta = 0.05$), as well as a colour and albedo similar to the Uranian satellites Umbriel and Oberon (Buratti, Goguen & Mosher 1997), mean that it may in fact have formed as a regular satellite of Neptune. For these reasons Bottke et al. (2010) did not model Neptunian irregulars.

Given these complications, it is perhaps unsurprising that our model for Neptune's irregulars needs very different $N_{\rm str}$ and α to the other giant planets, and is still a poor match. This difference is due to the slow collisional evolution, which predicts that $t_{\rm nleft}$ is a sizeable fraction of the Solar system age. With stranding occurring at such a late time, the subsequent evolution must be very rapid to deplete the population to that currently observed. With values for $N_{\rm str}$ and α more like the other three planets, our model predicts that Neptune would have two orders of magnitude more satellites (i.e. be at about the second to lowest curve). This discrepancy suggests that either the initial conditions for Neptune's swarm were quite different to the other planets, or as already proposed, the irregulars were affected by Nereid and Triton.

Based on these comparisons we conclude that our model provides a reasonable description of the collisional evolution of an irregular satellite swarm.

3.1.1 Current irregular dust levels

Given the ongoing collisional erosion of the irregular satellites, the presence of dust is inevitable. Here we make some estimates of the expected level of dust at each planet and relate them to a few relevant observations. Table 1 shows estimates for each planet based on $\sigma_{\rm tot}$ from the models of Fig. 2 at $t = 4.5 \times 10^9$ yr. That is, they are extrapolated using our size distribution and independent of the collisional model. We calculate the flux density and surface brightness at 1 and 100 μm (distributed uniformly over a disc of radius $0.5R_{\rm H}$ for simplicity), which roughly correspond to peaks in scattered light and thermal emission, respectively.

The predicted surface brightness levels are much fainter than the background in the ecliptic, but expected to vary on a similar scale (i.e. degrees). For comparison, the zodiacal background at 1.25 and 100 μm is about 0.4 and 9 MJy sr^{-1}, respectively (Kelsall et al. 1998). Detection of these clouds therefore requires accurate subtraction of this (and the Galactic and cosmic) background.

Of course dust that originates from irregular satellites has been detected already (see also Section 3.1.2). The largest of Saturn's ring is probably fed by material generated when interplanetary/circumplanetary grains impact the irregular satellite Phoebe (Verbiscer et al. 2009). At 24 μm, this ring has a surface brightness of ∼0.4 MJy sr^{-1}, which is less than 1 per cent of the (∼70 MJy sr^{-1}) zodiacal background. For comparison, our Saturn swarm model has ∼0.2 MJy sr^{-1} when spread over a $0.5R_{\rm H}$ radius disc.

To take a more detailed look at Saturn, Fig. 3 shows an example of what such a cloud might look like from Earth at 24 μm. In creating this image we have assumed that the dust follows the same orbits as the parent bodies, which have semimajor axes distributed between 0.04 and 0.16 au, and $e = 0.1$–0.6, based on fig. 7 of Nesvorný et al. (2007). The image would therefore look the same at any wavelength, though the total flux changes. The total flux in this 24-μm image is 320 Jy, which yields a peak surface brightness of

[1] Taken from http://www.dtm.ciw.edu/users/sheppard/satellites/ in 2010 March.

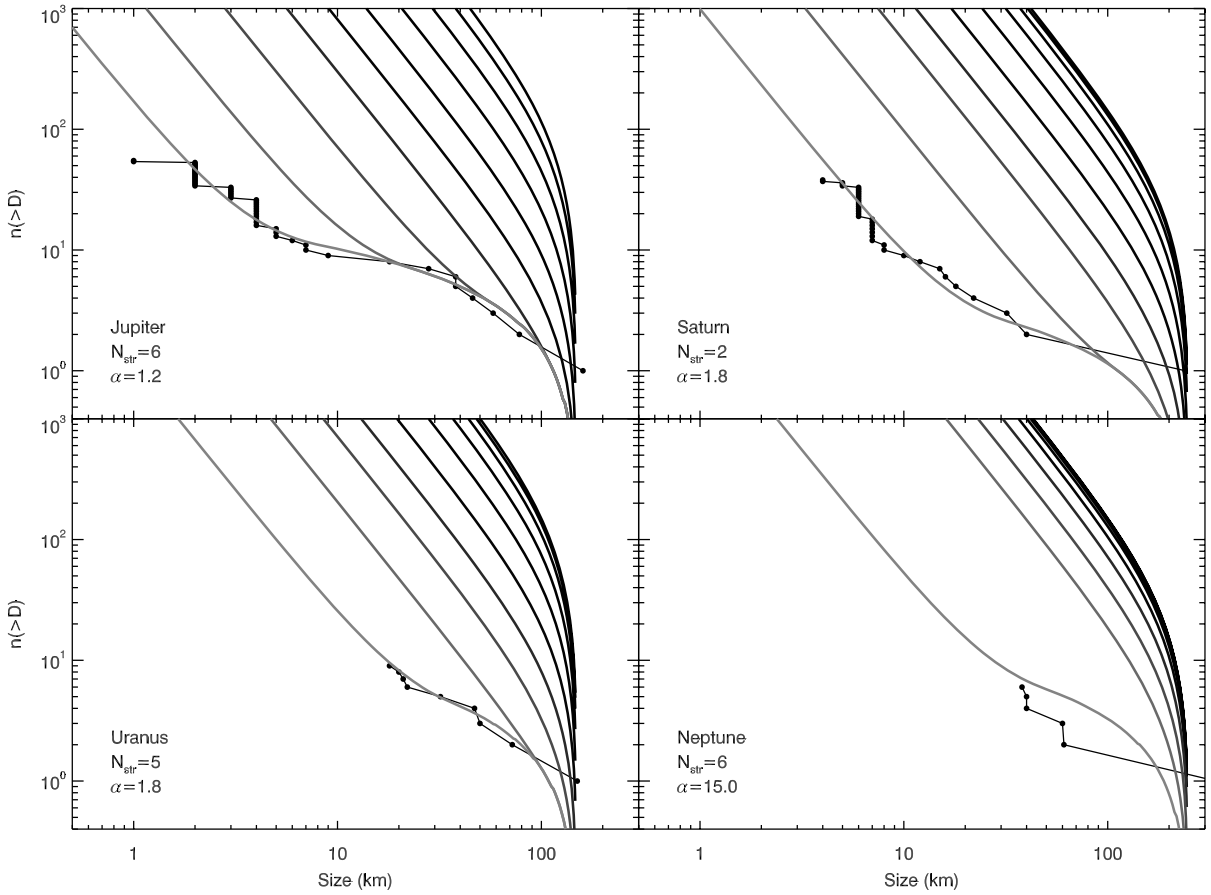

Figure 2. Comparison of our model (thick curves) with the Solar system's irregular satellites (dots and thin lines, prograde/retrograde populations have been added together). Curves show 10 logarithmically spaced times between 10^6 and 4.5×10^9 yr (dark to light). The legend indicates the planet and values for $N_{\rm str}$ and α. Model curves turn down at the largest sizes because they are derived from integrating $n(D)$ (and there are zero objects larger than $D_{\rm max}$).

around 1.25 MJy sr^{-1}. Therefore, the surface brightness predictions in Table 1 are a factor of ∼6 too low if grains follow their parent body orbits. The peak level is about three times brighter than Saturn's Phoebe ring.

Though the cloud is spread over a much larger region of sky, for it to have evaded detection thus far, particularly in the Phoebe ring observations, it appears that our prediction is at least a few times too high. However, Verbiscer et al. (2009) note an apparent trend of increasing surface brightness towards the Phoebe ring, which could be due to dust from irregular satellites. The narrow vertical range of their fig. 3 does not constrain our prediction, because the Phoebe ring could be sitting on top of a larger background.

Several uncertainties with our model may explain this apparently high prediction for the level of dust at Saturn, the most likely being that small departures from our assumed size distribution can lead to large differences in the predicted surface area in dust. In addition, we have assumed that grains follow the orbits of their parent bodies, but Krivov et al. (2002) show that the detailed cloud structure is more complex. Finally, Table 1 shows that grains around Saturn may be in the PR-dominated regime, which can alter the cloud structure and reduce the dust level.

Despite our apparent overprediction of the cloud surface brightness, detection of such a cloud is difficult, not least because it may be unexpected. The large extent means achieving sufficient coverage (ideally the entire Hill sphere) is expensive at infrared

Table 1. Irregular satellite model and dust properties at each planet. Estimates are based on the models in Fig. 2. The last six columns show estimates of the total dust cloud and planet flux density at opposition, and the surface brightness if the cloud were evenly spread over a disc with radius $0.5R_{\rm H}$, at 1 and 100 μm. The peak surface brightness is ∼6 times higher than that shown here *if* grains are distributed as in Fig. 3.

						$\sigma_{\rm tot}$				1 μm			100 μm	
Planet	$R_{\rm H}$ (au)	η	$t_{\rm col}$ (Myr)	$t_{\rm nleft}$ (Myr)	$D_{\rm min}$ (μm)	(10^{-9} au^2)	dens (km^{-3})	$\chi_{\rm PR}$	$F_{\rm dust}$ (Jy)	$F_{\rm pl}$ (Jy)	B (MJy sr^{-1})	$F_{\rm dust}$ (Jy)	$F_{\rm pl}$ (Jy)	B (MJy sr^{-1})
Jupiter	0.35	0.4	2.5	170	12	1.5	0.9	0.05	12.0	37 000	0.0034	960	450 000	0.26
Saturn	0.43	0.3	17.0	740	16	7.5	1.5	0.8	5.7	2400	0.0035	840	60 000	0.52
Uranus	0.46	0.2	14.0	1100	24	16.0	1.1	13.0	0.8	28	0.0016	210	950	0.45
Neptune	0.77	0.2	250.0	3500	23	36.0	0.6	13.0	0.3	4	0.0005	99	350	0.19

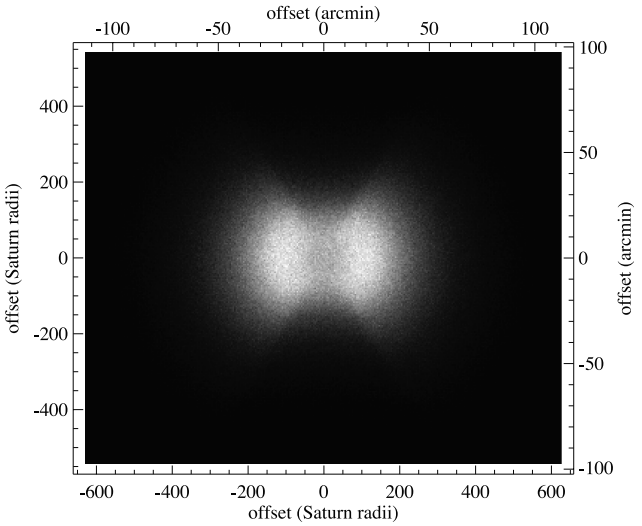

Figure 3. Simple model of Saturn's irregular satellite dust cloud at 24 μm. Grains follow parent body orbits. The total flux is 320 Jy and the scale is a linear stretch between 0 and 1.25 MJy sr^{-1}. Flux from Saturn itself is not included.

(IR) wavelengths (i.e. using space telescopes). Poorly characterized dust bands will also hinder background subtraction. At longer wavelengths, such as those covered by *Herschel* SPIRE, achieving decent coverage is easier, but the background comprises a combination of zodiacal and galactic light. Care would be needed to ensure good background subtraction that minimizes elongation effects.

Now looking at Jupiter, from Galileo dust detection data Krivov et al. (2002) derive a constant dust number density of $\sim$10 km^{-3} between 50 and 350$R_{\rm Jup}$. Krüger et al. (2010) report detection of micrometre-sized particles (and note the lack of detections of smaller particles) by the Galileo dust detector at a distance of approximately 350$R_{\rm Jup}$. Such large planetocentric distances are prime irregular satellite territory, being about half the Hill radius. These grains have previously been explained as ejecta from impacts of interplanetary dust on irregular satellites, which produce about the same level of dust as detected by Galileo (Krivov et al. 2002). The predicted space density of dust for Jupiter shown in Table 1 (again within 0.5$R_{\rm H}$) is at a lower level to the Galileo detections and may not have contributed to the measurements. However, the space density is an extremely strong function of $D_{\rm min}$. For example, if the minimum grain size in our model were 1 μm the predicted space density at Jupiter would be $\sim$100 km^{-3}. Therefore, Galileo observations constrain either the grain size in our model to be larger than a few micrometres, or the dust level to be lower than predicted if the minimum grain size is 1 μm.

3.1.2 *Fate of dust*

The best way to probe the small end of the irregular satellite size distribution is to detect the dust cloud directly. However, as noted above such an observation is difficult and there are complementary ways to detect irregular dust. One of the most interesting signatures exists on the surfaces of some regular satellites. The orbits of the smallest grains are strongly affected by radiation forces and may end up on a regular satellite. However, because this deposition is only one of several possible fates a grain may meet, an understanding of which is more likely (and when) is needed to make a strong connection between irregular dust and regular satellite surfaces. Our model is too simple to model irregular satellite evolution in the Solar system at a detailed level. However, we offer some order of magnitude arguments concerning the fate of dust, which highlight questions that should be asked by more detailed studies.

Like a circumstellar disc, a young circumplanetary swarm will most likely be collision dominated. At this stage grains are lost to the planet and interplanetary space. When the mass has been sufficiently depleted it becomes PR dominated and grains spiral in towards the planet. Grains destined to impact the planet, or pass nearby, may meet a third fate and be swept up by a regular satellite. Though only a small fraction of mass may be lost this way, it is important due to the visible effect of leading/trailing asymmetries on some tidally locked Solar system regular satellites, such as Callisto and Iapetus. These asymmetries are thought to arise from the higher accretion rate of retrograde dust by the leading hemisphere (e.g. Burns et al. 1996).

Though most grains in collision-dominated swarms either leave the Hill sphere or impact the planet due to radiation pressure, nearly all must pass through the regular satellite domain to do so. Grains destined to hit the planet or leave the Hill sphere do not reach their maximum eccentricity immediately, but instead make a number of pericentre passages first. The smallest grains with $\eta \sim 0.1$–0.5 only complete a few to a few tens of orbits before their eccentricities exceed unity (i.e. $t_{\rm per,grain} \sim \eta^{3/2} t_{\rm per,planet}$) and are therefore unlikely to encounter regular satellites before removal.

Grains in the small size range where e grows large enough to pass through the regular satellite region but not high enough to hit the planet or leave the Hill sphere have the best chance of colliding with regular satellites. However, this collision time must be shorter than the time for grains to collide with themselves. Taking Saturn as an example, at $\eta = 0.3$ with $\sigma_{\rm tot} = 7.5 \times 10^{-9}$ au^2 (Table 1) the current time for collisions between the smallest grains is predicted to be $\sim 10^6$ yr (equation 11). In the past the level of dust was much higher, and the grain–grain collision time correspondingly shorter.

Continuing with the Saturn example, grains with $0.83 > e > 0.997$ traverse the region between Iapetus' orbit and Saturn's surface near pericentre. This range corresponds to grains of sizes $16 > D > 25$ μm. The orbital period $t_{\rm per}$ is about 3 yr for these grains. Grains on coplanar orbits with Iapetus in this size range spend no more than 40 per cent of their lifetimes in Iapetus-crossing orbits. Iapetus occupies about 1/20 000th of its orbital torus, so coplanar grains have a collision time of $\sim 10^4$ orbits, or $\sim 10^4$ yr. This time is shorter than the grain–grain collision time, so coplanar grains in this size range appear likely to impact Iapetus.

However, grains on inclined orbits (i.e. the majority) only have a chance of hitting Iapetus if they cross Iapetus' plane at precisely the right radial distance. Iapetus is $\sim$1500 km across, compared to a semimajor axis of 3.5×10^6 km, so the chances of a plane-crossing particle encountering Iapetus *orbit* are roughly 1/2000. The chance of an inclined grain impacting Iapetus at each plane crossing is $\sim 10^{-7}$ per orbit, or an impact time of $\sim 10^7$ yr. The Iapetus collision time is therefore longer than the grain–grain collision time, so grains are more likely fragmented to smaller sizes first. Taking these numbers as representative, it appears that in collision-dominated swarms only a small fraction of grains, those within a small size range and on coplanar orbits, will impact regular satellites.

In contrast to the highly variable eccentricities of radiation pressure affected grains, PR drag causes grain orbits to collapse slowly. The chance of impacting a regular satellite is larger than for the more distant and eccentric radiation pressure induced orbits because grains orbit the planet more often as their semimajor axes shrink. For example, in the case of grains released from Phoebe,

collision with Iapetus is nearly guaranteed, with escapees destined to hit Hyperion or Titan (e.g. Burns et al. 1996).

Table 1 shows that dust from irregular satellites at Jupiter and Saturn is near or in the regime where PR drag dominates, but dust at Uranus and Neptune is not. If grains in PR-dominated swarms are more likely to impact regular satellites, this finding is consistent with the brightness asymmetries seen on Callisto and Iapetus, which are less marked (but still present) on the Uranian satellites.

While we suggest that the Jupiter and Saturn swarms are PR dominated now, they were collision dominated in the past. The possibility that collision-dominated swarms do not coat regular satellites as efficiently as PR-dominated ones therefore has implications for the interpretation of the brightness asymmetries. If the mass deposited on regular satellites is some fraction of the total mass lost, this fraction will increase as the swarm changes from collision to PR dominated. This change will affect the accretion history and more mass may be accreted at late stages if the difference in accretion efficiency is large enough.

Such interpretations may be too simple. For example, Iapetus' asymmetry is likely due to accretion of grains from Saturn's Phoebe ring, so is due to grains released from an individual irregular satellite rather than the cloud in general. In summary, further study of the fate of irregular satellite debris and potential observables requires consideration of the competition between collisions, PR drag and radiation pressure.

3.2 Irregular satellite clouds around exoplanets

While the dust produced by Solar system irregular satellites is at very low levels, these clouds were many times brighter at earlier times. In this section we explore the prospects for discovery of young extrasolar circumplanetary swarms. Like circumstellar debris discs, dust can be discovered in scattered light at optical wavelengths, and thermal emission at IR wavelengths. In the case of thermal emission, one may need to distinguish a planetary atmosphere from dust at a similar temperature via spectral features.

Because there is little a priori reason to choose any particular system configuration, we first model a particular system and then show which configurations are detectable in a more general sense. Clearly, to be detectable with current or near-future technology these systems must possess more dust than predicted for the Solar system's irregulars. Further, because the first objects discovered where sensitivity is an issue tend to be the most extreme (e.g. hot Jupiters), we do not restrict the initial swarm masses. While the initial conditions used by Bottke et al. (2010) based on the Nesvorný et al. (2007) simulations were at most a small fraction of the Moon's mass, extrasolar systems may have much more massive planetesimal belts from which to capture satellites, and mechanisms disfavoured for the Solar system may also operate. We discuss capture mechanisms further in Section 3.2.1.

In this section we first use the example of a Jupiter-mass planet orbiting an A5 star at 10 pc. We choose this spectral type of star simply because the only planets directly imaged to date orbit A-stars (Kalas et al. 2008; Marois et al. 2008; Lagrange et al. 2009). However, more massive stars do not necessarily produce brighter clouds because their high luminosity and mass increase $D_{\rm min}$ and change the collision rate. For the evolution of the planet, we use the non-irradiated Jupiter-mass ($Z = 0.02$) model from Baraffe, Chabrier & Barman (2008). We set $M_{\rm tot}(0) = 10M_{\rm moon}$, and $D_{\rm c} = 250$ km. Another choice is when the satellite swarm is captured (or otherwise formed) and evolution begins. For simplicity, we assume swarms start evolving at $t = 0$.

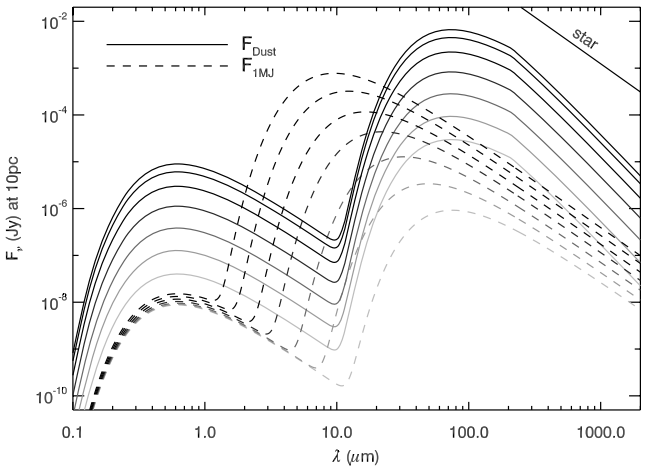

Figure 4. Example satellite swarm evolution around a $1\,M_{\rm Jup}$ extrasolar planet at 50 au around an A5 star between 10^7 (darkest curves) and 10^{10} yr (lightest curves, logarithmically spaced). A small part of the stellar spectrum is visible.

Fig. 4 shows the evolution of scattered light and thermal emission from both planet and dust cloud when the planet orbits at 50 au. The figure is drawn to highlight where the spectrum of each component moves over time (which is downwards in this plot). In scattered light, the dust cloud is initially much brighter than the planet, but decreases significantly as the satellite swarm collides and the dust level subsides. The planet shrinks somewhat as it cools, but in scattered light it is nearly constant. At early times the planet's thermal emission peaks slightly short of 10 μm, but moves to longer wavelengths over time. Thermal emission from the dust beyond 20 μm is brighter than the planet at all times. They become comparable at millimetre wavelengths at late times, which shows the importance of including grain emission inefficiencies.

This evolution therefore highlights wavelengths where irregular satellite clouds may be bright enough to be detected, both in absolute terms and relative to thermal and scattered emission from the planet. Unsurprisingly these lie at the thermal and stellar peaks. The key to detection lies with rejection of starlight, the same issue faced by those looking for planets or circumstellar debris discs (e.g. Beichman et al. 2007, 2010). This rejection is characterized by the star/planet contrast ratio that a particular instrument can detect, which is usually a function of angular separation between the two. While Fig. 4 shows that the youngest dust clouds may be bright enough for detection by the *Herschel* Photodetector Array Camera and Spectrometer (PACS) at 100 μm, the 7-arcsec resolution means that resolving these systems is difficult. We therefore focus on optical and near-/mid-IR wavelengths. We model the detectability of satellite swarms using simple approximations to published contrast ratios for several instruments.

For the *Hubble Space Telescope* Advanced Camera for Surveys (*HST* ACS) the contrast is based on actual roll-subtracted coronagraphic observations (Krist 2006) and we set an absolute detection limit of 0.1 μJy. We use predicted *James Webb Space Telescope* (*JWST*) Near-IR Camera (NIRCAM) 4.4 μm and Mid-IR Imager (MIRI) 11.4 contrast ratios from Beichman et al. (2010). For MIRI at 23 μm, we use the same contrast as at 11.4 μm for the same λ/D (i.e. at twice the separation, Boccaletti et al. 2004). The other difference from the three shorter MIRI coronagraph wavelengths is that only a Lyot stop is offered at 23 μm so the inner work-

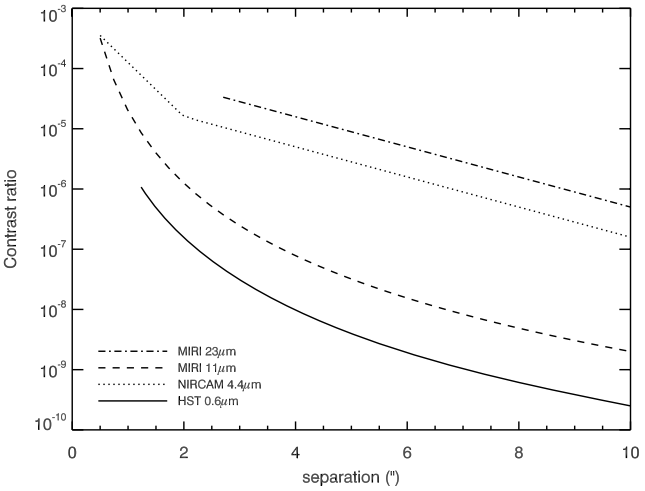

Figure 5. Adopted instrument contrast ratios.

ing angle is larger. We set absolute detection limits of 68 nJy, 2.5 μJy and 50 μJy for 4.4, 11.4 and 23 μm, respectively. These sensitivities are based on a 5σ detection in the difference of two 1-h exposures (Beichman et al. 2010). The contrast ratios are shown in Fig. 5.

Fig. 6 shows detectability contours for the Jupiter-mass planet at a range of possible planetary semimajor axes from our A5 star at 10 pc. For each instrument there are two curves; one for the swarm (black curve) and the other for the planet (grey curve). Swarms and planets in the space to the left of curves are detectable. The optimum planet semimajor axis for detection in scattered light for our chosen parameters (equation 16) is drawn as a dotted line.

Looking at the detection limits for each instrument individually, *HST* ACS detects scattered light from swarms across a wide range of the parameter space. As expected from equation (16), these swarms lie many tens of au from the star. Detectable swarms are at larger semimajor axes at later times as swarms around planets on closer orbits deplete. As expected from Fig. 4, Jupiter-mass planets are hard to detect with ACS at any separation (so there is no grey *HST* curve in Fig. 6).

JWST instruments are well suited to planet detection by design so that planets and swarms are detectable. The detection space for satellite swarms with NIRCAM covers semimajor axes greater than 70–120 au until a Gyr. Planets are detectable for a shorter time but to much closer separations. The difference is because the swarms are detected in scattered light, and planets from their thermal emission. At 11 μm MIRI does not detect the satellite swarms, but is ideal for detecting cooling planets, whose SEDs initially peak in or near this region. The detection region therefore covers even more age and separation space than at 4.4 μm.

At 23 μm MIRI detects the cooling planet, but also sees the Wein side of the dust cloud for systems up to 3×10^8 yr old. At this wavelength the visibility space for MIRI lies closer than suggested by $a_{\rm opt}$ because the dust clouds are hotter. With the exception of the vertical parts of curves (most notably at 23 μm), detectability in this example is set by contrast rather than absolute sensitivity.

Therefore, optical wavelengths appear to be the best place to detect extrasolar irregular satellite swarms around Jupiter-mass planets. Though detectable, at IR wavelengths the swarms are fainter than their host planets. Discovery then involves finding whether a (presumably previously known) planet has an optical or IR excess.

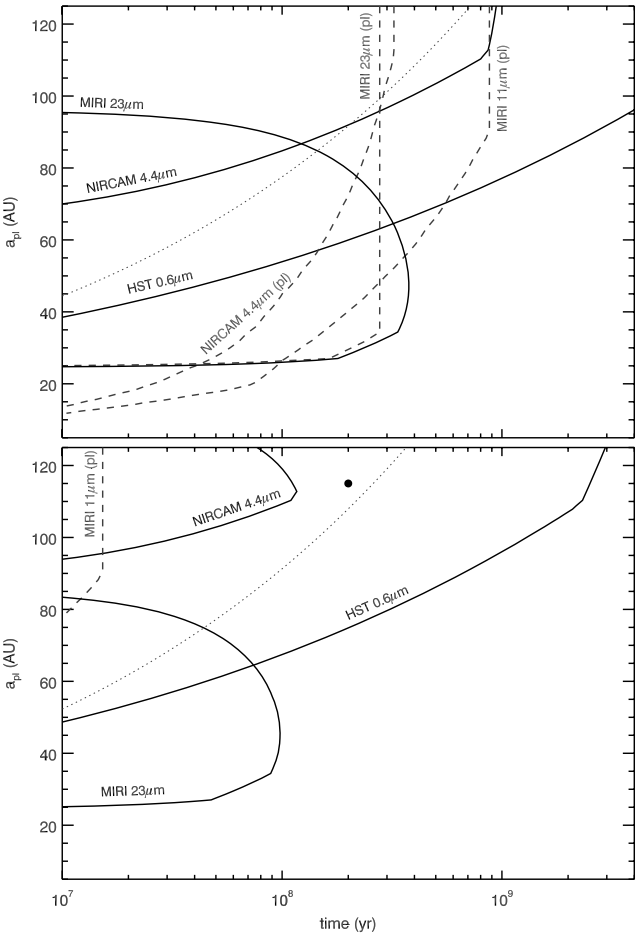

Figure 6. Regions in $a_{\rm pl}$ versus time space where satellite swarms (dark lines) and their host planets (grey dashed lines) can be detected with various instruments. All space to the left of each contour is detectable by that instrument (swarms and planets become fainter as they move towards right). The dashed line shows $a_{\rm opt}$ (equation 16). The top panel shows detection contours for a Jupiter-mass host, and the lower panel contours for a 20 $M_\oplus$ planet. Fomalhaut b is marked by a filled circle (see Section 3.3).

Over the parameter space in Fig. 6 we find that the largest excess ratio ($F_{\rm swarm}/F_\star$) is about 2 per cent at 100 μm and 3 per cent at 160 μm for a young swarm at 70–80 au. These excesses would not be detectable given current uncertainties with *Herschel* PACS photometry (Poglitsch et al. 2010) and planet atmospheres. Given the resolution, a sufficiently bright swarm detected this way would be classed as a circumstellar disc unless it were resolved at shorter wavelengths or orbiting a nearby star. Such a limitation may be overcome by high-resolution facilities such as the Atacama Large Millimeter Array (ALMA).

The predictions of Fig. 6 for all instruments can be extended in several directions. For the same star moved to 5 pc, swarms and planets at the same semimajor axis are more detectable because they lie at larger angular separations (where the instrument contrast ratio is better). Alternatively, a swarm around a planet at 10 arcsec separation around a star at 10 pc (planet at 100 au) is about equally detectable at optical/near-IR wavelengths when the star is at 5 pc (planet at 50 au), but more easily detectable at mid-IR wavelengths because the dust is hotter. For the same range of planet semimajor axes shown in Fig. 6, the same swarms become much harder to detect beyond a few tens of parsecs. The difficulty arises due to

the poorer contrast at smaller angular separations. If swarms are detectable beyond a few tens of parsecs, they must have of the order of an Earth mass or more in irregular satellites.

As the stellar mass decreases the detection space at optical/near-IR wavelengths increases due to the much lower stellar luminosity and consequently smaller minimum grain size. However, the clouds become undetectable in the mid-IR because the Wien side of their spectrum is too faint. The planets themselves become somewhat more detectable at IR wavelengths because their thermal emission is the same but the star is fainter (and we have assumed that all planetary luminosity is intrinsic).

Decreasing the planet mass can make for undetectable planets that host detectable swarms. The weak Hill radius dependence on planet mass means that collisional evolution is only somewhat slower (equation 7). In addition, the minimum grain size only depends weakly on planet mass (equation 9). To highlight the relative unimportance of planet mass on the evolution of satellite swarms, the lower panel of Fig. 6 shows detectability contours for a $20\,M_{\oplus}$ planet. The much fainter planet can only be detected at the earliest times with MIRI at 11 μm. Despite the 16 times decease in planet mass, the *HST* detection space is not much smaller than for the Jupiter-mass planet in the upper panel and still covers a wide range of orbits and ages. The detection space for NIRCAM at 4.4 μm is extremely small. At 23 μm, MIRI can still detect young swarms orbiting planets at 25–80 au from the star. This figure therefore shows that relatively low-mass planets that would otherwise be invisible can be detected thanks to the luminosity of their irregular satellite swarms. As we argue below for the specific case of Fomalhaut b, these swarms could be misidentified as more massive planets in the first instance.

3.2.1 Discussion

The ability of planets to capture irregular satellite swarms and reside at tens of au sets the likelihood that any one will be detected outside the Solar system. The wide range of planet masses about which swarms should be detectable means that how capture mechanisms and efficiencies change with planet mass is important. In particular, our prediction that swarms may be visible around relatively low-mass planets relies on the ability of these planets to capture swarms.

Giants forming by core accretion may capture irregulars passing through their primordial envelopes (e.g. Pollack et al. 1979), or by 'pull-down' during a phase of rapid growth (e.g. Heppenheimer & Porco 1977). Giants that form by gravitational instability may have analogous capture mechanisms. Jewitt & Haghighipour (2007) suggest that the similarity of the irregular populations at each planet argues against gas giant specific capture processes for the Solar system. However, Bottke et al. (2010) note that because the size distributions are a result of the collisional evolution, they cannot be used to constrain the capture mechanism.

Any planet may capture satellites during three- and n-body interactions (e.g. Colombo & Franklin 1971; Agnor & Hamilton 2006). Planets may capture satellites as they themselves interact in the presence of a planetesimal disc. Irregulars captured earlier by gas drag and pull-down are supplanted by those captured during planet–planet interactions. Such a scenario has been proposed for the origin of the Solar system's irregular satellites within the context of the Nice model (Nesvorný et al. 2007). The strength of this model is that it results in similar populations at each planet. In fact, in most model runs Nesvorný et al. (2007) find that Jupiter captures the least irregular satellites, because it undergoes the fewest encounters with other planets. The prospects for lower mass planets harbouring swarms of irregulars are therefore good.

Because we predict that swarms are most detectable around planets with large semimajor axes, the probability of detection depends on the ability of planets to form on, or move to, such orbits. The recent discovery of planets around HR 8799 (Marois et al. 2008) and Fomalhaut (Kalas et al. 2008) on such wide orbits is encouraging, not only due to their very existence, but because both stars also harbour planetesimal discs (Aumann 1984; Sadakane & Nishida 1986). Planets that originate on closer orbits must either scatter or migrate to such distances. If these systems have planetesimal discs, satellite capture during the scattering and migration process is likely. In such a scenario, the swarm's radial extent (η) cannot be much larger than half the Hill radius at the planetary semimajor axis where the swarm was captured (see also Section 3.3.3).

The presence or absence of irregular swarms within the context of discovered planetary systems should provide information about planet formation and evolution. For example, the Nesvorný et al. (2007) scenario requires the apparently specialized circumstances of a set of giant planets that are destabilized and interact in the presence of a planetesimal disc, but appears to be a robust way to capture irregulars. Such a scenario suggests that swarms may be discovered around planets in multiple systems that harbour debris discs (such as HR8799), especially those where scattering rather than migration has occurred. On the other hand, if capture of irregulars is associated with major planetesimal depletion events (as proposed by Nesvorný et al. 2007), systems with irregular swarms may tend *not* to have visible debris discs. If primordial irregulars are formed in all planetary systems, the properties of swarm-harbouring systems may be much more general. Therefore, it will be interesting to study the expected detection outcomes for a range of capture and formation mechanisms.

As noted previously, there are effects beyond the scope of our model that can affect our predictions. More theoretical effort should be made to understand which are important and how they affect our conclusions, in particular the effect of radiation forces on dust. Efforts should be made to detect dust clouds around Solar system giant planets, not only for their intrinsic interest, but to provide an empirical conversion between the mass in irregulars and the surface area in dust to calibrate our model predictions.

3.3 Fomalhaut b

An interesting application of our model is the planet orbiting the nearby star Fomalhaut, which also harbours a narrow ring-like circumstellar debris disc (Kalas et al. 2008). The planet was predicted to exist based on the elliptical shape of the debris ring (Kalas et al. 2005). Using the sharpness of the inner edge of the dust ring Quillen (2006) estimated the planet's orbit and a range of possible masses between Neptune's and Saturn's. In fact the structure of the ring continues to provide the most stringent constraints on the planet's mass, with more recent modelling providing an estimate of $<3\,M_{\rm Jup}$ if this planet is the 'sole sculptor' of the disc (Chiang et al. 2009).

Chiang et al. note that Fomalhaut b need not be the only sculptor of the belt. This point is in part made because the derived orbital velocity for Fomalhaut b is marginally inconsistent with that expected if the disc and planet are apsidally aligned. With Fomalhaut b only observed at two epochs, this inconsistency is a minor issue at present, but highlights the possibility that another planet may be partly responsible for the offset and truncation of the debris disc, and Fomalhaut b may be much less massive.

Another reason the mass is so poorly constrained at present is the lack of information about the planet's spectrum (Kalas et al. 2008). Thus far it has only been detected in two bands (0.6 and 0.8 μm), the first of which shows a factor of 3 change in brightness between measurements at two epochs separated by 2 yr. Neglecting the potential temporal evolution, these observations and the non-detections in other wavebands more closely resemble reflected starlight than thermal emission from a planet.

These observations lead Kalas et al. to suggest that the planet hosts a circumplanetary ring akin to Saturn's, which would extend to at least 20 Jupiter radii for an assumed albedo of 0.4 to recover the observed fluxes. If the emission is indeed scattered light from dust then the planet mass could be much less than 3 $M_{\rm Jup}$ (see also Arnold & Schneider 2004). In this section we suggest that the observed spectrum could indeed be reflected starlight, scattered off the dusty debris from a circumplanetary swarm of irregular satellites.

A convenient parameter space for this problem is the cloud semimajor axis in units of Hill radii (η) versus the planet mass ($M_{\rm pl}$). The observed quantities constrain η, because the planet is not resolved and the disc size is limited by optical depth. The planet mass will be constrained because a circumplanetary swarm that evolves like our model will not survive for the age of Fomalhaut over all parts of this parameter space. That is, for a swarm orbiting a ∼Jupiter-mass planet to be unresolved it would have small η, where it would be rapidly depleted to undetectable levels.

3.3.1 *Observational constraints*

The scattered light model in Kalas et al. (2008) requires a total cross-sectional area of dust:

$$\sigma_{\rm tot} = 5.8 \times 10^{-4}(0.08/Q)(1/\cos i)\,{\rm au}^2, \quad (24)$$

where we assume $Q = 0.08$ (see Section 2.4) and $\cos i$ is the factor that would be required if the geometry was that of a flat ring inclined at an angle of i to our line of sight, assumed to be ∼1 as the model of Section 2 is optically thin.

The light from Fomalhaut b looks unresolved so the dust must be confined to a region smaller than the Hubble point spread function (PSF) full width at half-maximum (FWHM) of 0.53 au. This region has an area of 0.22 au^2 so the geometrical optical depth of the dust could be as low as 2.6×10^{-3}. The resolution constraint means that $\eta r_{\rm H} \lesssim 0.53$ so $\eta < 0.6/M_{\rm pl}^{1/3}$, where we have allowed the dust cloud to be a factor of 2 larger in extent to account for the concentration of brightness closer to the planet seen in Fig. 3. As noted in Section 2, another constraint on η comes from the stability of circumplanetary orbits, $\eta < 0.5$.

Assuming that the dust is uniformly projected on the sky across $\pi(\eta r_{\rm H})^2$, its geometrical optical depth is $\tau = \sigma_{\rm tot}/(\pi(\eta r_{\rm H})^2)$. Since $\tau < 1$ the dust must be located at $\eta > 0.015/M_{\rm pl}^{1/3}$, which sets a lower limit on planet mass of $26 \times 10^{-6}\,{\rm M}_\oplus$ (i.e. six times less massive than Ceres and 80 times less massive than Pluto).

The above constraints, and the dynamical constraint of $M_{\rm pl} \lesssim 3M_{\rm Jup}$ from Chiang et al. (2009), are summarized by solid lines in Fig. 7. The satellite cloud may reside anywhere in the region enclosed by the solid lines, which spans over seven orders of magnitude in planet mass, and more than two in semimajor axis. These constraints are independent of our model (except perhaps the $0.5R_{\rm H}$ stability limit), and apply to any circumplanetary dust population. The $\sim 20R_{\rm Jup}$ disc proposed by Kalas et al. (2008) would lie in the lower right-hand region of the allowed space in Fig. 7.

Our aim is now to narrow this parameter space further. Because the collision rate depends strongly on η, satellite clouds appear more likely to survive for the age of Fomalhaut around planets much less massive than Jupiter.

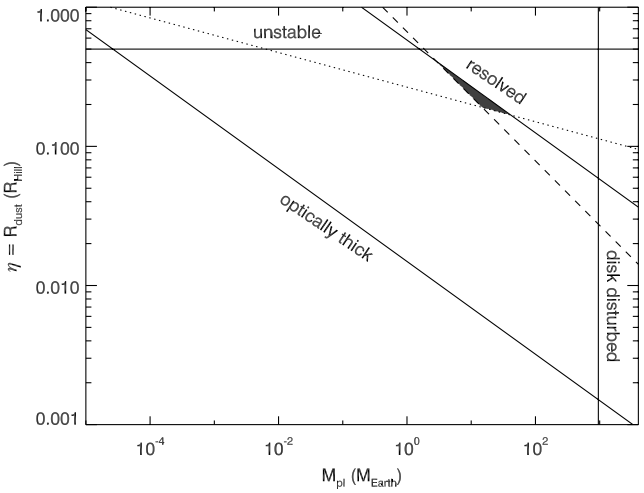

Figure 7. Constraints on the location of the proposed irregular circumplanetary swarm (solid lines) and loci indicating lower limits to allowed regions with our model (dotted and dashed lines). The region where a satellite swarm could survive for 200 Myr is enclosed by the dashed and dotted lines and the solid resolution line (shaded grey).

Currently, the typically quoted age of Fomalhaut is 200 Myr (Barrado y Navascues et al. 1997). However, recent work suggests that the age may be closer to 400 Myr (E. Mamajek, private communication). We adopt 200 Myr for the age of Fomalhaut, and note below the (small) effect doubling the age has on our model.

3.3.2 *Collisional constraints*

To find where satellite clouds survive for the age of Fomalhaut, we set the collision time ($1/R_{\rm cc}$) to 200 Myr (i.e. it is collision limited). It remains to substitute appropriate equations and estimates for parameters to reduce equation (7) to η as just a function of $M_{\rm pl}$. We use $M_\star = 2\,{\rm M}_\odot$, $L_\star = 21\,{\rm L}_\odot$, $\rho = 1500\,{\rm kg\ m}^{-3}$, $a_{\rm pl} = 115$ au and $D_{\rm c} = 500$ km. For comparison with Table 1, the Hill radii for Earth-, Neptune- and Jupiter-mass planets at this distance are 0.9, 2.3 and 6.2 au, respectively. We need to convert the observed $\sigma_{\rm tot}$ into $M_{\rm tot}$, for which we use equation (4). Because the effect of radiation pressure in setting $D_{\rm min}$ varies over the parameter space, we also need equation (9), which gives the mass as $M_{\rm tot} = 300\sigma_{\rm tot}\eta^{0.35}/M_{\rm pl}^{0.23}$. The mass in satellites is therefore about $0.06\,{\rm M}_\oplus$, or $5\,M_{\rm moon}$ for $\eta = 0.2$ and $M_{\rm pl} = 10\,{\rm M}_\oplus$, about 100–1000 times more massive than the initial conditions in the Bottke et al. (2010) models, and similar to Section 3.2. Such a low planet mass and higher stellar mass and luminosity result in $D_{\rm min} \sim 300$ μm, much higher than in the Solar system.

Substituting these parameters into equation (7) yields a locus for a collision-limited satellite swarm around Fomalhaut b that reproduces the observed $\sigma_{\rm tot}$. This locus $\eta = 0.27/M_{\rm pl}^{0.12}$ is shown in Fig. 7 as a dotted line. Because $D_{\rm c}$ may be smaller than our assumed 500 km, but based on the Solar system it is unlikely to be significantly larger, the space above this line is also allowed. Thus, the collision rate in concert with the resolution limit constrains the planet to have a maximum mass of about $100\,{\rm M}_\oplus$. The minimum mass is lower than the mass in satellites, and therefore physically unreasonable.

There is a further complication due to the large semimajor axis of Fomalhaut b, which results in low collision velocities $v_{\rm rel} = 68 M_{\rm pl}^{1/3}/\sqrt{\eta}$ (in m s^{-1}). A rough estimate of the largest object that can be destroyed by another object of the same size (and will therefore participate in the collisional evolution) can be derived from equation (2) and $X_{\rm c} = 1$ (Section 2.1), giving $D_{\rm c} \approx v_{\rm rel}^{1.6}/28$ (in km). This estimate yields another collision rate from equation (5) that applies when $D_{\rm c}$ is set by collision velocities

$$R_{\rm cc} = 114 \frac{M_{\rm tot} M_\star}{D_{\rm c}^{0.37} \rho (\eta a_{\rm pl})^3 M_{\rm pl}} \quad (25)$$

in yr^{-1}. This equation yields another locus in the η versus $M_{\rm pl}$ parameter space where the collision rate equals the system age of $\eta = 0.67/M_{\rm pl}^{0.46}$. This limit is shown in Fig. 7 as a dashed line. Though it would appear that the dotted line allows swarms with $D_{\rm c} = 500$ km to survive for planet masses right down to 10^{-3} M$_\oplus$, satellites this large cannot be destroyed at the low collision velocities around such low-mass planets. The dashed line is a lower limit because $D_{\rm c}$ could be smaller than the smallest object that can be destroyed. This limit therefore constrains the mass of Fomalhaut b to be more than a few Earth masses.

These two loci combined with the previous constraints map out a region of parameter space in Fig. 7 where a satellite swarm could survive for 200 Myr around Fomalhaut b. The planet is ∼2–100 M$_\oplus$ and the swarm lies at 0.1–0.4 Hill radii. The swarm mass is of the order of a few lunar masses (but varies with planet mass due to the changing $D_{\rm min}$, see above). The model predicts that a cloud of the observed luminosity could survive around a more massive planet, but this possibility is excluded because Fomalhaut b is unresolved. If Fomalhaut is 400-Myr old, the allowed parameter space is pushed even closer to the resolution limit, but does not change Fig. 7 significantly.

Fomalhaut b is marked in the lower panel of Fig. 6 (based only on age and orbit). As expected it lies within the region where swarms can be detected around a 20 M$_\oplus$ planet with *HST* and *JWST* NIRCAM.

Of course there is considerable uncertainty in both our model and the parameters it uses, but based on evidence that satellite swarms exist around Solar system planets, it at least suggests that circumplanetary swarms should be considered a possibility around extra solar planets. In the particular case of Fomalhaut b, the question of planet versus cloud can be resolved if new observations show that Fomalhaut b looks like a planetary atmosphere. In the event that scattered light cannot be ruled out, both circumplanetary rings (e.g. Kalas et al. 2008) and swarms remain plausible options.[2]

3.3.3 Discussion

While our model provides a possible explanation of Fomalhaut b's apparently blue spectrum, the provenance of such a configuration is unclear. Any scenario faces the difficulty of explaining how Fomalhaut b came to be at such a large distance from Fomalhaut itself. In a core accretion scenario, the planet presumably originates somewhere much closer to the star (e.g. Kennedy & Kenyon 2008), and somehow scatters or migrates to its current location. Because the Hill radius expands as the planet moves outwards, the satellite cloud would need to be captured while the planet was orbiting at least ∼20 au from the star for it to reside at $\eta > 0.1$ now. Capture of satellites might happen as Fomalhaut b scatters off other planets or migrates through a planetesimal disc on the way to its current location. The gas drag and pull-down mechanisms are unlikely to operate because our predicted mass of Fomalhaut b is insufficient for it to have a significant gaseous envelope.

Our predicted mass for Fomalhaut b is similar to or less than the mass of the main debris ring itself, which is ∼3–300 M$_\oplus$ (e.g. Wyatt & Dent 2002; Chiang et al. 2009). If the planet and ring masses are similar, it is unlikely that Fomalhaut b is responsible for truncating and imposing eccentricity on ring particles. Even if such a low-mass planet could reproduce the debris ring structure, the relatively small chaotic zone width may require that the planet lie closer to the ring than observed.

These issues do not necessarily pose a major problem for our model. Chiang et al. (2009) emphasize that Fomalhaut b may not actually be responsible for sculpting Fomalhaut's debris ring, in part because the planet's orbit is mildly inconsistent with the expected apsidal alignment of planet and ring particle orbits. However, an as-yet undiscovered object massive enough to sculpt the debris ring may have more serious implications for the stability of Fomalhaut b.

3.3.4 An alternative model

The current poor constraint on Fomalhaut b's orbit allows for other interesting possibilities. For example, its orbit may pass through the circumstellar ring. When the planet is within the ring, planetesimal impacts will generate a surrounding cloud of regolith, mantle and planetesimal fragments. The material may free-fall back to the planet as in the model of Wyatt & Dent (2002), or in a picture more like the formation of the Earth–Moon system or Kuiper belt binaries some fraction of the fragments can remain in bound orbits. If Fomalhaut b spends a non-negligible fraction of its orbit within the ring, possible if the planet orbit and ring are coplanar, this scenario provides a mechanism by which a population of objects bound to the planet may be built up and replenished, thus avoiding the requirement that the swarm must survive for the age of Fomalhaut to be observed. This way, even a small amount of mass launched into orbit each time the planet passes through the ring will build up over time, eventually reaching an equilibrium state where the bound mass is limited by collisions. Finding this equilibrium mass is then simply a matter of estimating the rate at which mass is added to the swarm and balancing it with the depletion rate. Because we expect the new satellites to be launched into near-isotropic orbits (i.e. to look like an irregular satellite swarm), the depletion rate is calculated from the model developed in Section 2.

The rate at which new satellites are added to the swarm is uncertain, so we adopt the approach taken by Wyatt & Dent (2002); an impacting planetesimal can at most launch its own mass of regolith and fragments to a significant distance from the planet. For impacts by large planetesimals these collisions may be more akin to the 'graze and capture' like scenario suggested for Pluto–Charon and Haumea's satellites (e.g. Canup 2005; Leinhardt, Marcus & Stewart 2010). Not all planetesimals accreted by the planet will launch their own mass in regolith and fragments away from the planet, nor will all launched material end up in orbit. We therefore expect that the accretion rate must be at least a few orders of magnitude higher than the loss rate for the existence of a cloud to be feasible.

[2] Patient observers may someday find photometric phase variations that would be caused by an optically thick circumplanetary disc (Arnold & Schneider 2004). However, other methods will likely become available before any variation can be found due to Fomalhaut b's $\sim 10^3$ yr period.

The gravitationally focused mass accretion rate can be calculated from

$$\dot{M}_{\rm acc} = \frac{M_{\rm ring}}{2\,\pi\, r_{\rm ring}^2\, {\rm d}r_{\rm ring}\, i_{\rm ring}}\, \pi R_{\rm pl}^2 \left(1 + \frac{v_{\rm esc}^2}{v_{\rm rel}^2}\right)\, v_{\rm rel}. \quad (26)$$

We assume $r_{\rm ring} = 141$ au, ${\rm d}r_{\rm ring} = 25$ au and $i_{\rm ring} = 0.026$ rad (using ring parameters from Kalas et al. 2005). We assume that the largest ring planetesimal is 500 km in diameter, which corresponds to a ring mass of about 75 $M_\oplus$ (Kalas et al. 2005), which is intermediate to the two cases considered in Wyatt & Dent (2002). The accretion rate of a 10 $M_\oplus$ planet with mass density 5000 kg m^{-3} is about 10^{-10} $M_\oplus$ yr^{-1}. Because the planet cannot spend all its time within the ring, we (arbitrarily) decrease the accretion rate by a factor of 2. The accretion rate would be much lower for a non-coplanar orbit, on which the planet would spend much less time (if any) within the ring. For a 10 $M_\oplus$ planet with mass density 5000 kg m^{-3}, and $v_{\rm rel} = 0.1\, v_{\rm K}$, equation (26) reduces to $2 \times 10^{-10}\, M_{\rm pl}^{4/3}$ $M_\oplus$ yr^{-1}.

The mass-loss rate is simply the current mass times the collision rate, which for a swarm in collisional equilibrium is also the inverse of the age (equation 14). More generally, using equations (2), (4) and (7), the mass-loss rate is

$$\dot{M}_{\rm loss} = 0.002 \left[\frac{\rho^{0.37} D_{\rm min}^{1.4} M_\star^{1.38} f_{v_{\rm rel}}^{2.27}}{M_{\rm pl}^{0.24} \left(a_{\rm pl}\, \eta\right)^{4.13}}\right] \sigma_{\rm tot}^2 \quad (27)$$

in $M_\oplus$ yr^{-1}. For our purposes here, the mass-loss rate is simply proportional to the square of the observed $\sigma_{\rm tot}$. That is, more mass is lost and at a faster rate for higher $\sigma_{\rm tot}$ or $M_{\rm tot}$. All parameters in the large parenthesis depend on the particular model assumptions. The rate is independent of the maximum planetesimal size $D_{\rm c}$. This independence can be viewed as due to our assumption of a self-similar size distribution, where the mass lost is the same in each logarithmic size bin. If the small end of the size distribution is fixed, larger $D_{\rm c}$ means more mass is available to be depleted ($M_{\rm tot}$), but collisions are less frequent and the largest objects are stronger.

Evaluating equation (27) for Fomalhaut b yields $\dot{M}_{\rm loss} = 5.5 \times 10^{-12} / \left(\eta^{3.4} M_{\rm pl}^{0.7}\right)$. Using the system age and a planet with a collision-limited swarm of mass 0.06 $M_\oplus$ from the previous section yields 3×10^{-10} $M_\oplus$ yr^{-1}, in agreement with the general expression for these parameters.

Using these two expressions for the mass accretion and loss rates and assuming an efficiency parameter $f_{\rm acc}$, a swarm will replenish as fast as it decays when $f_{\rm acc}\dot{M}_{\rm acc} = \dot{M}_{\rm loss}$. Fig. 8 shows where this equation is satisfied for different $f_{\rm acc}$. The line where the planet mass is too large to be consistent with the observed debris ring structure is omitted because a planet that passes through the ring will perturb it in a different way. This line would move to lower planet masses, and constrain this model. Another constraint on the upper planet mass would come from the mass above which a significant atmosphere is present, meaning that ring planetesimals are engulfed, rather than launch regolith and fragments. However, an atmosphere under periodic bombardment would be more prone to thermal escape due to heating.

The location of the lines in Fig. 8 suggest that replenishing a swarm from ring planetesimals is difficult. Even if a tenth of the accreted mass is launched into orbit the planet must be more than an Earth mass, which seems likely to significantly perturb the main ring on a time-scale less than 200 Myr. One way to lower the planet mass but maintain the accretion rate is if the planet is a binary. The cross-section for an interaction with a passing planetesimal scales with the binary semimajor axis and is therefore much larger than the physical size of either object (Hills 1975; Leonard 1989). Such an interaction does not guarantee a collision, but allows a greater chance of one for the same (total) planet mass. This may allow a binary to launch sufficient material into a swarm, yet also have a low enough mass to minimize dynamical perturbations to Fomalhaut's debris ring.

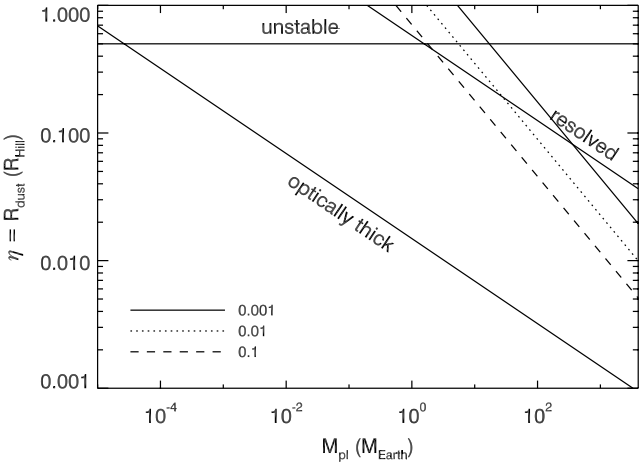

Figure 8. Lines in η versus $M_{\rm pl}$ parameter space where swarms can be replenished by accretion with different efficiency parameters $f_{\rm acc}$ (shown in legend). Though the 'disc-disturbed' limit is not drawn, we expect it to lie at a much lower planet mass than in Fig. 7.

4 SUMMARY AND CONCLUSIONS

The purpose of this paper is to derive a simple but general model for collisional evolution of irregular satellites and apply it to planets in the Solar system and elsewhere. Our model uses the particle-in-a-box formalism to describe the collision rate of the largest objects in a satellite swarm. A model size distribution allows the mass evolution to be converted to the surface area in dust and flux densities. Though the model can be developed further in many ways, a comparison with the Solar system irregular satellites shows that our model is reasonable.

Application of our model to the Solar system suggests that the bulk of grains may be lost to the planet or interplanetary space when the cloud is collision dominated. It may be that deposition on regular satellites only becomes important at later times when PR drag starts to dominate grain orbital evolution. Some level of dust must be present if the irregular satellites are still grinding down, which we suggest may be at detectable levels at any of the Solar system's outer planets.

A remarkable feature of irregular satellite swarm evolution is its relative insensitivity to planet mass. Swarms are nearly as bright and last nearly as long around Neptune-mass planets as around Jupiter-mass planets. In contrast to extrasolar planets, we find that satellite swarms are most visible at optical wavelengths around planets that orbit at many tens of au from their parent stars. Lower mass planets without gaseous envelopes cannot capture swarms by gas drag and pull-down, but dynamical mechanisms can still operate. There is an optimum distance for detection, which arises because swarms around planets on close orbits decay too rapidly and swarms around planets on distant orbits take too long to collide.

We propose a plausible model for a satellite swarm around Fomalhaut b. This model provides an alternative explanation for the spectrum of the planet being consistent with reflected starlight. In this picture Fomalhaut b is predicted to be ∼1–100 $M_\oplus$, further

illustrating that planet mass is relatively unimportant for the evolution of irregular satellite swarms. The allowed parameter space for the model lies very close to the resolution limit, so observations at higher resolution (FWHM < 0.53 au or 69 mas) should test our model. However, the order-of-magnitude spirit of our model and large uncertainties in model parameters that affect our allowed parameter space, such as satellite size, composition and strength, mean that the best way to test our model in this particular instance is to ascertain whether Fomalhaut b has the spectrum expected from a planet.

We briefly outline another possibility for a swarm around Fomalhaut b based on speculation that it could pass through Fomalhaut's large debris ring. In this case, a swarm of objects is maintained by regolith and fragments launched by planetesimal impacts. The ∼Earth mass planet required by such a model may be too high to avoid large (and unobserved) perturbations to the main ring. This issue may be somewhat alleviated if the planet is actually a binary, which would enhance the accretion rate and allow a lower total planet mass.

The prospects for detecting dust created by irregular satellite collisions, both in the Solar system and around planets orbiting other stars appear good. Along with further characterization of Fomalhaut b, the four populations of irregulars on our doorstep seem a good place to start.

ACKNOWLEDGMENTS

We are grateful to the Isaac Newton Institute for Mathematical Sciences in Cambridge where part of this work was carried out during the Dynamics of Discs and Planets research programme. We would also like to thank Paul Kalas for helpful comments and sharing preliminary observations of Fomalhaut b, Eric Mamajek for sharing his results on Fomalhaut's age and the referee Bill Bottke for a constructive review.

REFERENCES

Agnor C. B., Hamilton D. P., 2006, Nat, 441, 192
Arnold L., Schneider J., 2004, A&A, 420, 1153
Aumann H. H., 1984, BAAS, 16, 483
Baraffe I., Chabrier G., Barman T., 2008, A&A, 482, 315
Barrado y Navascues D., Stauffer J. R., Hartmann L., Balachandran S. C., 1997, ApJ, 475, 313
Beichman C. A., Fridlund M., Traub W. A., Stapelfeldt K. R., Quirrenbach A., Seager S., 2007, Protostars and Planets V. Univ. Arizona Press, Tucson
Beichman C. A. et al., 2010, PASP, 122, 162
Benz W., Asphaug E., 1994, Icarus, 107, 98
Benz W., Asphaug E., 1999, Icarus, 142, 5
Boccaletti A., Riaud P., Baudoz P., Baudrand J., Reess J., Rouan D., 2004, in Aime C., Soummer R., eds, EAS Publ. Ser. Vol. 12, Coronagraphy with JWST in the Thermal IR. EDP Sciences, Paris, p. 195
Bottke W. F., Nolan M. C., Greenberg R., Kolvoord R. A., 1994, Icarus, 107, 255
Bottke W. F., Nesvorný D., Vokrouhlický D., Morbidelli A., 2010, AJ, 139, 994
Buratti B. J., Goguen J. D., Mosher J. A., 1997, Icarus, 126, 225
Burns J. A., Lamy P. L., Soter S., 1979, Icarus, 40, 1
Burns J. A., Hamilton D. P., Mignard F., Soter S., 1996, in Gustafson B. A. S., Hanner M. S., eds, ASP Conf. Ser. Vol. 104, IAU Colloq. 150, Physics, Chemistry, and Dynamics of Interplanetary Dust. Astron. Soc. Pac., San Francisco, p. 179
Campo Bagatin A., Cellino A., Davis D. R., Farinella P., Paolicchi P., 1994, Planet. Space Sci., 42, 1079
Canup R. M., 2005, Sci, 307, 546
Carruba V., Burns J. A., Nicholson P. D., Gladman B. J., 2002, Icarus, 158, 434
Chiang E., Kite E., Kalas P., Graham J. R., Clampin M., 2009, ApJ, 693, 734
Collier Cameron A., Horne K., Penny A., Leigh C., 2002, MNRAS, 330, 187
Colombo G., Franklin F. A., 1971, Icarus, 15, 186
Cox A. N., 2000, Allen's Astrophysical Quantities. Springer Verlag, Berlin
Ćuk M., Gladman B. J., 2005, ApJ, 626, L113
Dohnanyi J. S., 1969, J. Geophys. Res., 74, 2531
Dominik C., Decin G., 2003, ApJ, 598, 626
Durda D. D., Greenberg R., Jedicke R., 1998, Icarus, 135, 431
Fernández Y. R., Jewitt D., Ziffer J. E., 2009, AJ, 138, 240
Gladman B. et al., 2001, Nat, 412, 163
Goldreich P., Murray N., Longaretti P. Y., Banfield D., 1989, Sci, 245, 500
Gomes R., Levison H. F., Tsiganis K., Morbidelli A., 2005, Nat, 435, 466
Gustafson B. A. S., 1994, Annu. Rev. Earth Planet. Sci., 22, 553
Heng K., Tremaine S., 2010, MNRAS, 401, 867
Heppenheimer T. A., Porco C., 1977, Icarus, 30, 385
Hills J. G., 1975, AJ, 80, 809
Holman M. J. et al., 2004, Nat, 430, 865
Horanyi M., 1996, ARA&A, 34, 383
Jewitt D., Haghighipour N., 2007, ARA&A, 45, 261
Kalas P., Graham J. R., Clampin M., 2005, Nat, 435, 1067
Kalas P. et al., 2008, Sci, 322, 1345
Kelsall T. et al., 1998, ApJ, 508, 44
Kennedy G. M., Kenyon S. J., 2008, ApJ, 673, 502
Kessler D. J., 1981, Icarus, 48, 39
Krist J., 2006, in Traub W., ed., Coronograph Workshop 2006 (NASA). p. 83
Krivov A. V., Wardinski I., Spahn F., Krüger H., Grün E., 2002, Icarus, 157, 436
Krüger H. et al., 2010, Planet. Space Sci., 58, 965
Lagrange A. et al., 2009, A&A, 493, L21
Leinhardt Z. M., Marcus R. A., Stewart S. T., 2010, ApJ, 714, 1789
Leonard P. J. T., 1989, AJ, 98, 217
Levison H. F., Bottke W. F., Gounelle M., Morbidelli A., Nesvorný D., Tsiganis K., 2009, Nat, 460, 364
Marois C., Macintosh B., Barman T., Zuckerman B., Song I., Patience J., Lafrenière D., Doyon R., 2008, Sci, 322, 1348
Morbidelli A., Levison H. F., Tsiganis K., Gomes R., 2005, Nat, 435, 462
Mueller M., Grav T., Trilling D., Stansberry J., Sykes M., 2008, BAAS, 40, 511
Nesvorný D., Alvarellos J. L. A., Dones L., Levison H. F., 2003, AJ, 126, 398
Nesvorný D., Vokrouhlický D., Morbidelli A., 2007, AJ, 133, 1962
O'Brien D. P., Greenberg R., 2003, Icarus, 164, 334
Poglitsch A. et al., 2010, A&A, 518, L2+
Pollack J. B., Burns J. A., Tauber M. E., 1979, Icarus, 37, 587
Quillen A. C., 2006, MNRAS, 372, L14
Sadakane K., Nishida M., 1986, PASP, 98, 685
Shen Y., Tremaine S., 2008, AJ, 136, 2453
Sheppard S. S., Jewitt D. C., 2003, Nat, 423, 261
Sheppard S. S., Jewitt D., Kleyna J., 2005, AJ, 129, 518
Stansberry J., Grundy W., Brown M., Cruikshank D., Spencer J., Trilling D., Margot J., 2008, in Barucci M. A., Boehnhardt H., Cruikshank D. P., Morbidelli A., eds, Physical Properties of Kuiper Belt and Centaur Objects: Constraints from the Spitzer Space Telescope, The Solar System Beyond Neptune. Univ. Arizona Press, Tucson, p. 161
Stewart S. T., Leinhardt Z. M., 2009, ApJ, 691, L133
Tsiganis K., Gomes R., Morbidelli A., Levison H. F., 2005, Nat, 435, 459
Verbiscer A. J., Skrutskie M. F., Hamilton D. P., 2009, Nat, 461, 1098

Wyatt M. C., 2008a, in Mann I., Nakamura A. M., Mukai T., eds, Lecture Notes in Physics 758, Small Bodies in Planetary Systems. Springer Verlag, Berlin, p. 329
Wyatt M. C., 2008b, ARA&A, 46, 339
Wyatt M. C., Dent W. R. F., 2002, MNRAS, 334, 589
Wyatt M. C., Dermott S. F., Telesco C. M., Fisher R. S., Grogan K., Holmes E. K., Piña R. K., 1999, ApJ, 527, 918
Wyatt M. C., Smith R., Greaves J. S., Beichman C. A., Bryden G., Lisse C. M., 2007, ApJ, 658, 569
Wyatt M. C., Booth M., Payne M. J., Churcher L. J., 2010, MNRAS, 402, 657

This paper has been typeset from a TEX/LATEX file prepared by the author.