The Multifaceted Planetesimal Formation Process

Anders Johansen
Lund University

Jürgen Blum
Technische Universität Braunschweig

Hidekazu Tanaka
Hokkaido University

Chris Ormel
University of California, Berkeley

Martin Bizzarro
Copenhagen University

Hans Rickman
Uppsala University

Polish Academy of Sciences Space Research Center, Warsaw

Accumulation of dust and ice particles into planetesimals is an important step in the planet formation process. Planetesimals are the seeds of both terrestrial planets and the solid cores of gas and ice giants forming by core accretion. Left-over planetesimals in the form of asteroids, trans-Neptunian objects and comets provide a unique record of the physical conditions in the solar nebula. Debris from planetesimal collisions around other stars signposts that the planetesimal formation process, and hence planet formation, is ubiquitous in the Galaxy. The planetesimal formation stage extends from micrometer-sized dust and ice to bodies which can undergo run-away accretion. The latter ranges in size from 1 km to 1000 km, dependent on the planetesimal eccentricity excited by turbulent gas density fluctuations. Particles face many barriers during this growth, arising mainly from inefficient sticking, fragmentation and radial drift. Two promising growth pathways are mass transfer, where small aggregates transfer up to 50% of their mass in high-speed collisions with much larger targets, and fluffy growth, where aggregate cross sections and sticking probabilities are enhanced by a low internal density. A wide range of particle sizes, from mm to 10 m, concentrate in the turbulent gas flow. Overdense filaments fragment gravitationally into bound particle clumps, with most mass entering planetesimals of contracted radii from 100 to 500 km, depending on local disc properties. We propose a hybrid model for planetesimal formation where particle growth starts unaided by self-gravity but later proceeds inside gravitationally collapsing pebble clumps to form planetesimals with a wide range of sizes.

1. INTRODUCTION

Most stars are born surrounded by a thin protoplanetary disc with a characteristic mass between 0.01% and 10% of the mass of the central star (Andrews and Williams, 2005). Planetesimal formation takes place as the embedded dust and ice particles collide and grow to ever larger bodies. Tiny particles collide gently due to Brownian motion, while larger aggregates achieve higher and higher collision speeds as they gradually decouple from the smallest eddies of the turbulent gas flow and thus no longer inherit the incompressibility of the gas flow (Voelk et al., 1980). The gas disc is partially pressure-supported in the radial direction, causing particles of sizes from centimeters to meters to drift towards the star (Whipple, 1972; Weidenschilling, 1977a). Drift speeds depend on the particle size and hence differently sized particles experience high-speed collisions.

The growth from dust and ice grains to planetesimals – the seeds of both terrestrial planets as well as the cores of gas giants and ice giants – is an important step towards the assembly of a planetary system. A planetesimal can be defined as a body which is held together by self-gravity rather than material strength, corresponding to minimum sizes of order 100–1,000 meters (Benz, 2000). The planetesimal formation stage, on the other hand, must extend to sizes...
where the escape speed of the largest bodies exceeds the random motion of the planetesimals, to enter the next stage of gravity-driven collisions. The random speed of planetesimals and preplanetesimals (the latter can be roughly defined as bodies larger than 10 meters in size) is excited mainly by the stochastic gravitational pull from density fluctuations in the turbulent gas; under nominal turbulence conditions the largest planetesimals need to reach sizes as large as 1,000 km to start run-away growth (Ida et al., 2008). In disc regions with less vigorous turbulence, planetesimal sizes between 1 and 10 kilometers may suffice to achieve significant gravitational focusing (Gressel et al., 2012).

Planetesimals must grow to run-away sizes despite bouncing and disruptive collisions (Blum and Wurm, 2000). Planetesimals must also grow rapidly – radial drift timescales of cm-m-sized particles are as short as a few hundred orbits. Despite these difficulties there is ample evidence from cosmochemistry that large planetesimals formed in the solar nebula within a few million years (Kleine et al., 2004; Baker et al., 2005), early enough to melt and differentiate by decay of short-lived radionuclides.

The time since the last Protostars & Planets review on planetesimal formation (Dominik et al., 2007) has seen a large expansion in the complexity and realism of planetesimal formation studies. In this review we focus on three areas in which major progress has been obtained, namely (i) the identification of a bouncing barrier at mm sizes and the possibility of growth by mass transfer in high-speed collisions (Wurm et al., 2005; Johansen et al., 2008; Güttler et al., 2010; Zsom et al., 2010; Windmark et al., 2012b), (ii) numerical simulations of collisions between highly porous ice aggregates which can grow past the bouncing and radial drift barriers due to efficient sticking and cross sections that are greatly enhanced by a low internal density (Wada et al., 2008, 2009; Seizinger and Kley, 2013), and (iii) hydrodynamical and magnetohydrodynamical simulations which have identified a number of mechanisms for concentrating particles in the turbulent gas flow of protoplanetary discs, followed by a gravitational fragmentation of the overdense filaments to form planetesimals with characteristic radii larger than 100 km (Rice et al., 2004; Johansen et al., 2007; Lyra et al., 2008a, 2009; Johansen et al., 2009a; Bai and Stone, 2010a; Raettig et al., 2013).

The review is laid out as follows. In section 2 we give an overview of the observed properties of planetesimal belts around the Sun and other stars. In section 3 we review the physics of planetesimal formation and derive the planetesimal sizes necessary to progress towards planetary systems. Section 4 concerns particle concentration and the density environment in which planetesimals form, while the following section 5 describes laboratory experiments to determine the outcome of collisions and methods for solving the coagulation equation. In section 6 we discuss the bouncing barrier for silicate dust and the possibility of mass transfer in high-speed collisions. Section 7 discusses computer simulations of highly porous ice aggregates whose low internal density can lead to high growth rates. Finally in section 8 we integrate the knowledge gained over recent years to propose a unified model for planetesimal formation involving both coagulation, particle concentration and self-gravity.

2. PLANETESIMAL BELTS

The properties of observed planetesimal belts provide important constraints on the planetesimal formation process. In this section we review the main properties of the asteroid belt and trans-Neptunians population as well as debris discs around other stars.

2.1. Asteroids

The asteroid belt is a collection of mainly rocky bodies residing between the orbits of Mars and Jupiter. Asteroid orbits have high eccentricities (\(e \sim 0.1\)) and inclinations, excited by resonances with Jupiter and also a speculated population of embedded super-Ceres-sized planetary embryos which were later dynamically depleted from the asteroid belt (Wetherill, 1992). The high relative speeds imply that the asteroid belt is in a highly erosive regime where collisions lead to fragmentation rather than to growth.

Asteroids range in diameters from \(D \approx 1000\) km (Ceres) down to the detection limit at sub-km sizes (Gladman et al., 2009). The asteroid size distribution can be parameterised in terms of the cumulative size distribution \(N_\geq(D)\) or the differential size distribution \(dN_\geq/dD\). The cumulative size distribution of the largest asteroids resembles a broken power law \(N_\geq \propto D^{-p}\), with a steep power law index \(p \approx 3.5\) for asteroids above 100 km diameter and a shallower power law index \(p \approx 1.8\) below this knee (Bottke et al., 2005).

Asteroids are divided into a number of classes based on their spectral reflectivity. The main classes are the moderate-albedo S-types which dominate the region from 2.1 to 2.5 AU and the low-albedo C-types which dominate regions from 2.5 AU to 3.2 AU (Gradie and Tedesco, 1982; Bus and Binzel, 2002). If these classes represent distinct formation events, then their spatial separation can be used to constrain the degree of radial mixing by torques from turbulent density fluctuations in the solar nebula (Nelson and Gressel, 2010). Another important class of asteroids is the M-type of which some are believed to be the metallic cores remaining from differentiated asteroids (Rivkin et al., 2000). Two of the largest asteroids in the asteroid belt, Ceres and Vesta, are known to be differentiated (from the shape and measured gravitational moments, respectively, Thomas et al., 2005; Russell et al., 2012).

Asteroids are remnants of solar system planetesimals that have undergone substantial dynamical depletion and collisional erosion. Dynamical evolution models can be used to link their current size distribution to the primordial birth size distribution. Bottke et al. (2005) suggested, based on the observed knee in the size distribution at sizes around 100 km, that asteroids with diameters above \(D \approx 120\) km are primordial and that their steep size distribution reflects their formation sizes, while smaller asteroids are fragments
of collisions between their larger counterparts. Morbidelli et al. (2009) tested a number of birth size distributions for the asteroid belt, based on either classical coagulation models starting with km-sized planetesimals or turbulent concentration models predicting birth sizes in the range between 100 and 1000 km (Johansen et al., 2007; Cuzzi et al., 2008), and confirmed that the best match to the current size distribution is that the asteroids formed big, in that the asteroids above 120 km in diameter are depleted dynamically but have maintained the shape of the primordial size distribution. On the other hand, Weidenschilling (2011) found that an initial population of 100-meter-sized planetesimals can reproduce the current observed size distribution of the asteroids, including the knee at 100 km. However, gap formation around large planetesimals and trapping of small planetesimals in resonances is not included in this particle-in-a-box approach (Levison et al., 2010).

2.2. Meteorites

Direct information about the earliest stages of planet formation in the solar nebula can be obtained from meteorites that record the interior structure of planetesimals in the asteroid belt. The reader is referred to the chapter by Davis et al. for more details on the connection between cosmochemistry and planet formation. Meteorites are broadly characterized as either primitive or differentiated (Krot et al., 2003). Primitive meteorites (chondrites) are fragments of parent bodies that did not undergo melting and differentiation and, therefore, contain pristine samples of the early solar system. In contrast, differentiated meteorites are fragments of parent bodies that underwent planetesimal-scale melting and differentiation to form a core, mantle and crust.

The oldest material to crystallise in the solar nebula is represented in chondrites as mm- to cm-sized calcium-aluminium-rich inclusions (CAIs) and ferromagnesian sili- cate spherules (chondrules) of typical sizes from 0.1 mm to 1 mm (Fig. 1). The chondrules originated in an unidentified thermal processing mechanism (or several mechanisms) which melted pre-existing nebula solids (e.g. Deusch and Connolly, 2002; Ciesla et al., 2004); alternatively chondrules can arise from the crystallisation of splash ejecta from planetesimal collisions (Sanders and Taylor, 2005; Krot et al., 2005).

The majority of CAIs formed as fine-grained condensates from an $^{16}$O-rich gas of approximately solar composition in a region with high ambient temperature ($>1300$ K) and low total pressures ($\sim 10^{-4}$ bar). This environment may have existed in the innermost part of the solar nebula during the early stage of its evolution characterized by high mass accretion rates ($\sim 10^{-5} M_\odot \text{ yr}^{-1}$) to the proto-Sun (D’Alessio et al., 2005). Formation of CAIs near the proto-Sun is also indicated by the presence in these objects of short-lived radionuclide $^{10}$Be formed by solar energetic particle irradiation (McKeegan et al., 2000). In contrast, most chondrules represent coalesced $^{16}$O-poor dust aggregates that were subsequently rapidly melted and cooled in lower-temperature regions ($<1000$ K) and higher ambient vapor pressures ($\geq 10^{-3}$ bar) than CAIs, resulting in igneous porphyritic textures we observe today (Scott, 2007). Judging by their sheer abundance in chondrite meteorites, the formation of chondrules must reflect a common process that operated in the early solar system. Using the assumption-free uranium-corrected Pb-Pb dating method, Connelly et al. (2012) recently showed that CAIs define a brief formation interval corresponding to an age of 4567.30 ± 0.16 Myr, whereas chondrule ages range from 4567.32 ± 0.42 to 4564.71 ± 0.30 Myr. These data indicate that chondrule formation started contemporaneously with CAIs and lasted ~3 Myr.

A consequence of accretion of planetesimals on Myr time-scales or less is the incorporation of heat-generating short-lived radioisotopes such as $^{26}$Al, resulting in widespread melting, differentiation and extensive volcanic activity. Both long-lived and short-lived radioisotope chronometers have been applied to study the timescale of asteroidal differentiation. Of particular interest are the $^{26}$Al-$^{26}$Mg and $^{182}$Hf-$^{182}$W decay systems, with half-lives of 0.73 Myr and 9 Myr, respectively. Al and Mg are refractory and lithophile elements and thus remain together in the mantle after core formation. In contrast, the different geochemical behavior of Hf and W during melting results in W being preferentially partitioned in the core and Hf being partitioned into the silicate mantle and crust. Therefore, the short-lived $^{182}$Hf-$^{182}$W system is useful to study the timescales of core formation in asteroids, as well as in planets, which can be used to constrain planetesimal formation models.

Eucrite and angrite meteorites are the two most common groups of basaltic meteorites, believed to be derived from the mantles of differentiated parent bodies. The HED (howardite-eucrite-diogenite) meteorite clan provides our best samples of any differentiated asteroid and could come from the 500-km-diameter asteroid 4 Vesta (Binzel and Xu,
ogy of chondrule formation of Connelly et al. (2009) requires that most chondrite parent bodies accreted constrained by the ages of their youngest chondrules, which timing of accretion of chondritic parent bodies can be con-

larly to chondrite meteorites, the main original constituents the epoch of chondrule formation. This suggests that, simi-

larly, the Mg isotope composition of olivines within pallasite mete-
orites - a type of stony-iron meteorites composed of cm-
sized olivine crystals set in a iron-nickel matrix - suggests silicate differentiation within 1.5 Myr of solar system for-

mation (Baker et al., 2012). Chemical and isotopic diversity of iron meteorites show that these have sampled approxi-
mately 75 distinct parent bodies (Goldstein et al., 2009). The $^{182}$Hf-$^{182}$W systematics of some magmatic iron mete-
orites require the accretion and differentiation of their par-

ent bodies to have occurred within less than 1 Myr of solar system formation (Kleine et al., 2009).

An important implication of the revised absolute chronol-
ogy of chondrule formation of Connelly et al. (2012), which is not based on short-lived radionuclides for which homo-

genicity in the solar nebula has to be assumed, is that the ac-
cretion and differentiation of planetesimals occurred during the epoch of chondrule formation. This suggests that, simi-

larly to chondrite meteorites, the main original constituents of early-accreted asteroids may have been chondrules. The timing of accretion of chondritic parent bodies can be con-

strained by the ages of their youngest chondrules, which requires that most chondrite parent bodies accreted >2 Myr

after solar system formation. Therefore, chondrites may represent samples of asteroidal bodies that formed after the accretion of differentiated asteroids, at a time when the lev-

els of $^{26}$Al were low enough to prevent significant heating and melting.

2.3. Trans-Neptunian objects

The trans-Neptunian objects are a collection of rocky/icy objects beyond the orbit of Neptune (Luu and Jewitt, 2002). Trans-Neptunian objects are categorised into several dy-

namical classes, the most important for planetesimal for-
mation being the classical Kuiper belt objects, the scattered disc objects and the related centaurs which have orbits that cross the orbits of one or more of the giant planets. The classical Kuiper belt objects do not approach Neptune at any point in their orbits and could represent the pristine pop-

ulation of icy planetesimals in the outer solar nebula. The so-called cold component of the classical Kuiper belt has a large fraction (at least 30%) of binaries of similar size (Noll et al., 2008) which strongly limits the amount of collisional grinding that these bodies can have undergone (Nesvorný et al., 2011). The scattered disc objects on the other hand can have large semi-major axes, but they all have perihelia close to Neptune’s orbit. The centaurs move between the gi-

ant planets and are believed to be the source of short period comets (see below).

The largest trans-Neptunian objects are much larger than the largest asteroids, with Pluto, Haumea, Makemake and Eris defined as dwarf planets of 1.5-3 times the diame-
ter of Ceres (Brown, 2008). Nevertheless the size distribu-
tion of trans-Neptunian objects shows similarities with the asteroid belt, with a steep power law above a knee at around $D\sim100$ km (Fuentes and Holman, 2008). The turn-

over at the knee implies that there are fewer intermediate-

mass planetesimals than expected from an extrapolation from larger sizes – this has been dubbed the “Missing Intermediate-Sized Planetesimals” problem (Sheppard and Trujillo, 2010; Shankman et al., 2013) and suggests that the characteristic planetesimal birth size was $\sim100$ km. The accretion ages of Kuiper belt objects are not known, in contrast to the differentiated asteroids where the inclusion of large amounts of $^{26}$Al requires early accretion. While the largest Kuiper belt objects are likely differentiated into a rocky core and an icy mantle – this is clear e.g. for Haumea which is the dense remnant of a differentiated body (Ragouzine and Brown, 2009) – differentiation could be due to long-lived radionuclides such as U, Th and $^{40}$K and hence happen over much longer time-scales ($\sim100$ Myr) than the $\sim$Myr time-scale characteristic of asteroid differ-

entiation by $^{26}$Al decay (McKinnon et al., 2008).

2.4. Comets

Comets are typically km-sized volatile-rich bodies which enter the inner solar system. They bring with them a wealth of information about the conditions during the planetesimal formation epoch in the outer solar system. Out-

Fig. 2.— Surface and interior structure of the 500-km-diameter asteroid 4 Vesta, obtained by the NASA Dawn satellite. The iron core is 110 km in radius, surrounded by a silicate mantle (green) and a basaltic crust (gray). Image credit: NASA/JPL-Caltech
gassing provides knowledge about the compositions and heating histories of icy planetesimals, and the volumes and masses of comet nuclei can be used to derive their density which can be compared to the density expected from planetesimal formation models. Comet nuclei are (or are related to) icy planetesimals, i.e. planetesimals that formed beyond the snow line in a formation zone extending from roughly 15 to 30 AU from the Sun – at least in the framework of the so-called Nice Model where the giant planets form in a compact configuration between 5 and 12 AU and later migrate by planetesimal scattering to their current positions (Levison et al., 2011). This zone is wide enough to allow for the appearance of chemical zoning, but such a zoning would likely not be reflected in any separation between the short- and long-period comet source regions.

Comets have to be considered together with the Centaurs and trans-Neptunians, because these are dynamically related to the Jupiter Family comets, and most of the data we have on comets come from observations and modeling of Jupiter Family comets. The emergent picture is a very wide size spectrum going from sub-km to 10^3 km in diameter. However, the slope of the size distribution is difficult to establish, since small objects are unobservable in the outer populations, and large ones are lacking in the Jupiter Family. Generally, the measurements of masses and sizes of comet nuclei are consistent with a rather narrow range of densities at about 0.5 g/cm^3 (Weissman et al., 2004; Davidsson et al., 2007). Most of the determinations use the non-gravitational force due to asymmetric outgassing. Assuming a composition that is roughly a 50-50 mix of ice and refractories (Greenberg and Hage, 1990) the porosity comes out as roughly 2/3. But the mass determinations only concern the bulk mass, so one cannot distinguish between meso-scale-porosity (rubble piles) and small-scale porosity (pebble piles). Estimates of comet tensile strengths have generally been extremely low, as expected for porous bodies. For comet Shoemaker-Levy 9, modeling of its tidal breakup led to (non-zero) values so low that the object was described as a strengthless rubble pile (Asphaug and Benz, 1996). The non-tidal splittings often observed for other comets appear to be so gentle that, again, an essentially zero strength has been inferred (Sekanina, 1997).

The current results on volatile composition of comets have not indicated any difference between the Jupiter Family comets – thought to probe the scattered disc – and the Halley-type and long-period comets, which should come from the Oort Cloud (A’Hearn et al., 2012). From a dynamical point of view, this result is rather expected, since both these source populations have likely emerged from the same ultimate source, a disc of icy planetesimals extending beyond the giant planet zone in the early Solar System (Brasser and Morbidelli, 2013).

The origin of comets is closely related to the issue of the chemical pristineness of comets. If the total mass of the planetesimal disc was dominated by the largest planetesimals, several hundreds of km in size, then the km-sized comet nuclei that we are familiar with could arise from collisional grinding of the large planetesimals. Nevertheless there are severe problems with this idea. One is that comet nuclei contain extremely volatile species, like the S_2 molecule (Bockelée-Morvan et al., 2004), which would hardly survive the heating caused by a disruptive impact on the parent body. The alternative that tidal disruptions of large planetesimals at mutual close encounters may lead to large numbers of smaller objects seems more viable, but this too suffers from the second pristineness problem, namely, the geologic evolution expected within a large planetesimal due to short-lived radio nuclei (Prialnik et al., 2004). If the large planetesimal thus becomes chemically differentiated, there seems to be no way to form the observed comet nuclei by breaking it up – no matter which mechanism we invoke. The pristine nature of comets is thus consistent with the formation of small, km-scale cometesimals in the outer part of the solar nebula, avoiding due to their small size melting and differentiation due to release of short-lived radionuclides. Another possibility is that comets accreted from material comprising an early formed 26Al-free component (Olsen et al., 2013), which would have prevented differentiation of their parent bodies. The recent chronology of 26Al-poor inclusions found in primitive meteorites indicate that this class of objects formed coeval with CAIs (which record the canonical 26Al/27Al value of 5 × 10^{-5}). Holst et al., (2013). This provides evidence for the existence of 26Al-free material during the earliest stages of the protoplanetary disc. The discovery of refractory material akin to CAIs in the Stardust samples collected from Comet 81P/Wild 2 show efficient outward transport of material from the hot inner disc regions to cooler environments far from the Sun during the epoch of CAI formation (Brownlee et al., 2006). Efficient outward transport during the earliest stages of solar system evolution would have resulted in the delivery of a significant fraction of the 26Al-poor material to the accretion region of cometary bodies. Analysis of the Coci refractory particle returned from comet 81P/Wild 2 did not show detectable 26Al at the time of its crystallization (Matzel et al., 2010) which, although speculative, is consistent with the presence an early-formed 26Al-free component in comets.

2.5. Debris discs

Planetesimal belts around other stars show their presence through the infrared emission of the dust produced in collisions (Wyatt, 2008). The spectral energy distribution of the dust emission reveals the orbital distance of the planetesimal belt. Warm debris discs resembling the asteroid belt in the Solar System are common around young stars of around 50 Myr age (at least 30%), but their occurrence falls to a few percent within 100–1,000 Myr (Siegler et al., 2007). Thus planetesimal formation appears to be ubiquitous around the Sun as well as around other stars. This is in agreement with results from the Kepler mission that flat planetary systems exist around a high fraction of solar-type stars (Lissauer et al., 2011; Tremaine and Dong, 2012; Jo-
where $v$ is the particle velocity, $u$ is the gas velocity at the position of the particle and $\tau_f$ is the friction time which contains all the physics of the interaction of the particle with the gas flow (Whipple, 1972; Weidenschilling, 1977a). The friction time can be divided into different regimes, depending on the mean free path of the gas molecules, $\lambda$, and the speed of the particle relative to the gas, $\delta v = |v - u|$. The Epstein regime is valid when the particle size is smaller than the mean free path. The flux of impinging molecules is set in this regime by their thermal motion and the friction time is independent of the relative speed,

$$\tau_f = \frac{R \rho_\star}{c_s \rho_g} \ (3)$$

Here $R$ is the radius of the particle, assumed to be spherical. The other parameters are the material density $\rho_\star$, the gas sound speed $c_s$ and the gas density $\rho_g$. Particles with sizes above $9/4$ times the mean free path of the molecules enter the Stokes regime (see Whipple, 1972, and references therein), with

$$\tau_f = \frac{R \rho_\star}{c_s \rho_g} \frac{4}{9} \lambda . \ (4)$$

Here the friction time is proportional to the squared radius and independent of gas density, since $\lambda$ is inversely proportional to the gas density. The flow Reynolds number past the particle, $Re = (2R\delta v)/\nu$, determines the further transition to drag regimes that are non-linear in the relative speed, with the kinematic viscosity given by $\nu = (1/2)c_s\lambda$. At unity flow Reynolds number the drag transitions to an intermediate regime with the friction time proportional to $(\delta v)^{-0.4}$. Above $Re = 800$ the drag force finally becomes quadratic in the relative velocity, with friction time

$$\tau_f = \frac{6R \rho_\star}{(\delta v) \rho_g} \ (5)$$

Following the descriptions above, the step-wise transition from Epstein drag to fully quadratic drag happens in the optically thin minimum mass solar nebula (MMSN, Hayashi, 1981), with power-law index $-1.5$ for the surface density (Weidenschilling, 1977b) and $-0.5$ for the temperature, at particle sizes

$$R_1 = \frac{9 \lambda}{4} \approx 3.2 \text{ cm} \left( \frac{r}{\text{AU}} \right)^{2.75} \ (6)$$

$$R_2 = \frac{\nu}{2(\delta v)} \approx 6.6 \text{ cm} \left( \frac{r}{\text{AU}} \right)^{2.5} \ (7)$$

$$R_3 = \frac{800\nu}{2(\delta v)} \approx 52.8 \text{ m} \left( \frac{r}{\text{AU}} \right)^{2.5} \ (8)$$

The friction time can be divided into different regimes, depending on the mean free path of the gas molecules, $\lambda$, and the speed of the particle relative to the gas, $\delta v = |v - u|$. The Epstein regime is valid when the particle size is smaller than the mean free path. The flux of impinging molecules is set in this regime by their thermal motion and the friction time is independent of the relative speed, with $\tau_f = \frac{R \rho_\star}{c_s \rho_g} \ (3)$.

Here $R$ is the radius of the particle, assumed to be spherical. The other parameters are the material density $\rho_\star$, the gas sound speed $c_s$ and the gas density $\rho_g$. Particles with sizes above $9/4$ times the mean free path of the molecules enter the Stokes regime (see Whipple, 1972, and references therein), with $\tau_f = \frac{R \rho_\star}{c_s \rho_g} \frac{4}{9} \lambda . \ (4)$

Here the friction time is proportional to the squared radius and independent of gas density, since $\lambda$ is inversely proportional to the gas density. The flow Reynolds number past the particle, $Re = (2R\delta v)/\nu$, determines the further transition to drag regimes that are non-linear in the relative speed, with the kinematic viscosity given by $\nu = (1/2)c_s\lambda$. At unity flow Reynolds number the drag transitions to an intermediate regime with the friction time proportional to $(\delta v)^{-0.4}$. Above $Re = 800$ the drag force finally becomes quadratic in the relative velocity, with friction time $\tau_f = \frac{6R \rho_\star}{(\delta v) \rho_g} \ (5)$.

Following the descriptions above, the step-wise transition from Epstein drag to fully quadratic drag happens in the optically thin minimum mass solar nebula (MMSN, Hayashi, 1981), with power-law index $-1.5$ for the surface density (Weidenschilling, 1977b) and $-0.5$ for the temperature, at particle sizes $R_1 = \frac{9 \lambda}{4} \approx 3.2 \text{ cm} \left( \frac{r}{\text{AU}} \right)^{2.75} \ (6)$, $R_2 = \frac{\nu}{2(\delta v)} \approx 6.6 \text{ cm} \left( \frac{r}{\text{AU}} \right)^{2.5} \ (7)$, $R_3 = \frac{800\nu}{2(\delta v)} \approx 52.8 \text{ m} \left( \frac{r}{\text{AU}} \right)^{2.5} \ (8)$.
A natural dimensionless parameter to construct from the friction time is the Stokes number $St = \Omega \tau_f$, with $\Omega$ denoting the Keplerian frequency at the given orbital distance. The inverse Keplerian frequency is the natural reference time-scale for a range of range of physical effects in protoplanetary discs, hence the Stokes number determines (i) turbulent collision speeds, (ii) sedimentation, (iii) radial and azimuthal particle drift, (iv) concentration in pressure bumps and vortices and (v) concentration by streaming instabilities. Fig. 3 shows the particle size corresponding to a range of Stokes numbers, for both the nominal density of $1 \text{ g/cm}^3$ and extremely fluffy particles with an internal density of $10^{-5} \text{ g/cm}^3$ (which could be reached when ice aggregates collide, see section 7).

### 3.2. Radial drift

Protoplanetary discs are slightly pressure-supported in the radial direction, due to gradients in both mid-plane density and temperature. This leads to sub-Keplerian motion of the gas, $v_{\text{gas}} = v_K - \Delta v$ with $v_K$ denoting the Keplerian speed and the sub-Keplerian velocity difference $\Delta v$ defined as (Nakagawa et al., 1986)

$$\Delta v \equiv \eta v_K = -\frac{1}{2} \left( \frac{H}{r} \right)^2 \frac{\partial \ln P}{\partial \ln r} v_K. \tag{8}$$

In the MMSN the aspect ratio $H/r$ rises proportional to $r^{1/4}$ and the logarithmic pressure gradient is $\partial \ln P / \partial \ln r = -3.25$ in the mid-plane. This gives a sub-Keplerian speed which is constant $\Delta v = 53 \text{ m/s}$, independent of orbital distance. Radiative transfer models with temperature and density dependent dust opacities yield disc aspect ratios $H/r$ with complicated dependency on $r$ and thus a sub-Keplerian motion which depends on $r$ (e.g. Bell et al., 1997). Nevertheless, a sub-Keplerian speed of $\sim 50 \text{ m/s}$ can be used as the nominal value for a wide range of protoplanetary disc models.

The drag force on the embedded particles leads to particle drift in the radial and azimuthal directions (Whipple, 1972; Weidenschilling, 1977a)

$$v_r = -\frac{\Delta v}{St + St^2}, \tag{9}$$

$$v_\phi = v_K - \frac{\Delta v}{1 + St^2}. \tag{10}$$

These equations give the drift speed directly in the Epstein and Stokes regimes. In the non-linear and quadratic drag regimes, where the Stokes number depends on the relative speed, the equations can be solved using an iterative method to find consistent $v_r$, $v_\phi$ and $St$.

The azimuthal drift peaks at $v_\phi = v_K - \Delta v$ for the smallest Stokes numbers where the particles are carried passively with the sub-Keplerian gas. The radial drift peaks at unity Stokes number where particles spiral in towards the star at $v_r = -\Delta v$. The radial drift of smaller particles is slowed down by friction with the gas, while particles larger than Stokes number unity react to the perturbing gas drag by entering mildly eccentric orbits with low radial drift.

Radial drift puts requirements on the particle growth at Stokes number around unity (generally from $St = 0.1$ to $St = 10$) to occur within time-scale $t_{\text{drift}} \sim r/\Delta v \sim 100-1000$ orbits (Brauer et al., 2007, 2008a), depending on the location in the protoplanetary disc. However, three considerations soften the at first glance very negative impact of the radial drift. Firstly, the ultimate fate of drifting particles is not to fall into the star, but rather to sublimate at evaporation fronts (or snow lines). This can lead to pile up of material around evaporation fronts (Cuzzi and Zahnle, 2004) and to particle growth by condensation of water vapour onto existing ice particles (Stevenson and Lunine, 1988; Ros and Johansen, 2013). Secondly, the radial drift flow of particles is linearly unstable to streaming instabilities (Youdin and Goodman, 2005), which can lead to particle concentration in dense filaments and planetesimal formation by gravitational fragmentation of the filaments (Johansen et al., 2009a; Bai and Stone, 2010a). Thirdly, very fluffy particles with low internal density reach unity Stokes number, where the radial drift is highest, in the Stokes drag regime (Okuzumi et al., 2012). In this regime the Stokes number, which determines radial drift, increases as the square of the particle size and hence growth to “safe” Stokes numbers with low radial drift is much faster than for compact particles. These possibilities are discussed more in the following sections.

### 3.3. Collision speeds

Drag from the turbulent gas excites both large-scale random motion of particles as well as collisions (small-scale random motion). The two are distinct because particles may have very different velocity vectors at large separations, but these will become increasingly aligned as the particles approach each other due to the incompressible nature of the gas flow. Particles decouple from turbulent eddies with turn-over times shorter than the friction time, and the gradual decoupling from the smallest eddies results in crossing particle trajectories. This decoupling can be interpreted as singularities in the particle dynamics (so-called “caustics”, Gustavsson and Mehlig, 2011). Caustics in turn give rise to collisions as small particles enter regions of intense clustering (clustering is discussed further in section 4.1.1). The random motion is similar to the turbulent speed of the gas for small particles, but particles experience decreased random motion as they grow to Stokes numbers above unity.

The problem when calculating the turbulent collision speed is that at close separations (when the particles are about to collide) they interact with the same eddies, which
causes their motions to become highly correlated. The framework set out by Voelk et al. (1980) which employs a Langevin approach, is still widely used, and Ormel and Cuzzi (2007) provided closed-form analytical approximations to their results (but see Pan and Padoan, 2010, for a criticism of the simplifications made in the Völk model).

The closed-form expressions of Ormel and Cuzzi (2007) require numerical solution of a single algebraic equation for each collided particle pair (defined by their friction times). With knowledge of the properties of the turbulence, particularly the turbulent rms speed and the frequency of the smallest and the largest eddies, the collision speeds can then be calculated at all locations in the disc.

Another important contribution to turbulent collision speeds is the gravitational pull from turbulent gas density fluctuations. The eccentricity of a preplanetesimal increases as a random walk due to uncorrelated gravitational kicks from the turbulent density field (Laughlin et al., 2004; Nelson and Papaloizou, 2004). The eccentricity would grow unbounded with time as $e \propto t^{1/2}$ in absence of dissipation. Equating the eccentricity excitation time-scale with the time-scale for damping by tidal interaction with the gas disc (from Tanaka and Ward, 2004), aerodynamic gas drag, and inelastic collisions with other particles, Ida et al. (2008) provide parameterisations for the equilibrium eccentricity as a function of particle mass and protoplanetary disc properties. The resulting collision speeds dominate over the contributions from the direct drag from the turbulent gas at sizes above approximately 10 meters.

Ida et al. (2008) adopt the nomenclature of Ogihara et al. (2007) for the eccentricity evolution, where a dimensionless parameter $\gamma$ determines the proportionality between the eccentricity and $t^{1/2}$. The parameter $\gamma$ is expected to scale with the density fluctuations $\delta\rho/\rho$ but can be directly calibrated with turbulence simulations. The shearing box simulations by Yang et al. (2009) of turbulence caused by the magnetorotational instability (Balbus and Hawley, 1991) suggest that $\delta\rho/\rho \propto \sqrt{\alpha}$, where $\alpha$ is the dimensionless measure of the turbulent viscosity (Shakura and Sunyaev, 1973). In their nominal ideal-MHD turbulence model with $\alpha \approx 0.01$, Yang et al. (2012) find $\gamma \approx 6 \times 10^{-4}$. This leads to an approximate expression for $\gamma$ as a function of the strength of the turbulence, $\gamma \approx 0.006\sqrt{\alpha}$. The resulting eccentricities are in broad agreement with

Fig. 4.— The collision speed, in meters per second, of two particles of size $a_1$ and $a_2$, with contributions from Brownian motion, differential radial and azimuthal drift, and gas turbulence. The upper panels show collision speeds for $\alpha = 10^{-2}$ and the lower panels show collision speeds for $\alpha = 10^{-4}$. The gravitational pull from turbulent gas density fluctuations is included in the right panels. The red line marks the transition from dominant excitation by direct drag to dominant excitation by turbulent density fluctuations. The “oasis” of low collision speeds for particles above 10 meters vanishes when including eccentricity pumping by turbulent density fluctuations.
the resistive magnetohydrodynamics simulations of Gressel et al. (2012) where the turbulent viscosity from Reynolds stresses fall below $\alpha = 10^{-4}$ in the mid-plane.

In Fig. 4 we show the collision speeds of two particles of sizes from $\mu$m to 100 km, in a figure similar to Fig. 3 of the classical Protostars and Planets III review by Weidenschilling and Cuzzi (1993). We take into account the Brownian motion, the differential drift, and the gas turbulence (from the closed-form expressions of Ormel and Cuzzi, 2007). We consider an MMSN model at $r = 1$ AU and $\alpha = 10^{-2}$ in the upper panels and $\alpha = 10^{-4}$ in the lower panels. The gravitational pull from turbulent density fluctuations driven by the magnetorotational instability is included in the right panels. The collision speed approaches 100 m/s for meter-sized boulders when $\alpha = 10^{-2}$ and 30 m/s when $\alpha = 10^{-4}$, due to the combined effect of the drag from the turbulent gas and differential drift. The collision speed of large, equal-sized particles (with $St > 1$) would drop as

$$\delta v = c_s \sqrt{\frac{\alpha}{St}}$$  \hspace{1cm} (11)$$

in absence of turbulent density fluctuations (for a discussion of this high-St regime, see Ormel and Cuzzi, 2007). This region has been considered a safe haven for planetesimals after crossing the ridges around unity Stokes number (Weidenschilling and Cuzzi, 1993). However, the oases vanishes when including the gravitational pull from turbulent gas density fluctuations, and instead the collision speeds continue to rise towards larger bodies, as they are damped less and less by gas drag (right panels of Fig. 4).

The collision speed of equal-sized particles is explored further in Fig. 5. Here we show results both of a model with fully developed MRI turbulence and a more advanced turbulent stirring model which includes the effect of a dead zone in the mid-plane, where the ionisation degree is too low for the gas to couple with the magnetic field (Fleming and Stone, 2003; Oishi et al., 2007), with weak stirring by density waves travelling from the active surface layers (Okuzumi and Ormel, 2013; Ormel and Okuzumi, 2013).

The collision speeds rise with size until peaking at unity Stokes number with collision speed approximately $\sqrt{\alpha c_s}$. The subsequent decoupling from the gas leads to a decline by a factor ten, followed by an increase due to the turbulent density fluctuations starting at sizes of approximately 10 – 100 meters. The planetesimal formation stage can be defined as the growth until the gravitational cross section of the largest bodies is significantly larger than their geometric cross section, a transition which happens when the escape speed of the largest bodies approach their random speed. In Fig. 5 we indicate also the escape speed as a function of the size of the preplanetesimal. The transition to run-away accretion can only start at 1000 km bodies in a disc with nominal turbulence $\alpha = 10^{-2}$ (Ida et al., 2008). Lower values of $\alpha$, e.g. in regions of the disc where the ionisation degree is too low for the magnetorotational instability, lead to smaller values for the planetesimal size needed for run-away accretion, namely 1–10 km.

Fig. 5.— The collision speed of equal-sized particles as a function of their size, from dust to planetesimals, based on the stirring model of Ormel and Okuzumi (2013). The collision speeds of small particles (below approximately 10 m) are excited mainly by direct gas drag, while the gravitational pull from turbulent gas density fluctuations dominates for larger particles. The transition from relative-speed-dominated to escape-speed-dominated can be used to define the end of the planetesimal formation stage (dashed line). The planetesimal sizes that must be reached for run-away accretion range from near 1000 km in highly turbulent discs ($\alpha = 10^{-2}$) to around 1–10 km in disc regions with an extended dead zone that is stirred by sheared density waves from the active regions ($\alpha = 2 \times 10^{-4}$).
3.4. Sedimentation

While small micrometer-sized grains follow the gas density tightly, larger particles gradually decouple from the gas flow and sediment towards the mid-plane. An equilibrium mid-plane layer is formed when the turbulent diffusion of the particles balances the sedimentation. The Stokes number introduced in section 3.1 also controls sedimentation (because the gravity towards the mid-plane is proportional to $\Omega^2$). Balance between sedimentation and turbulent diffusion (Dubrulle et al., 1995) yields a mid-plane layer thickness $H_p$, relative to the gas scale-height $H$, of

$$\frac{H_p}{H} = \sqrt{\frac{\delta}{\text{St} + \delta}}.$$  \hspace{1cm} (12)

Here $\delta$ is a measure of the diffusion coefficient $D = \delta c_p H$, similar to the standard definition of $\alpha$ for turbulent viscosity (Shakura and Sunyaev, 1973). In general $\delta \approx \alpha$ in turbulence driven by the magnetorotational instability (Johansen and Klahr, 2005; Turner et al., 2006), but this equality may be invalid if accretion is driven for example by disc winds (Blandford and Payne, 1982; Bai and Stone, 2013).

Equation (12) was derived assuming particles to fall towards the mid-plane at their terminal velocity $v_z = -\pi \Omega^2 z$. Particles with Stokes number larger than unity do not reach terminal velocity before arriving at the mid-plane and hence undergo oscillations which are damped by drag and excited by turbulence. Nevertheless Carballido et al. (2006) showed that equation (12) is in fact valid for all values of the Stokes number. Youdin and Lithwick (2007) interpreted this as a cancellation between the increased sedimentation time of oscillating particles and their decreased reaction to the turbulent motion of the gas.

The particle density in the equilibrium mid-plane layer is

$$\rho_p = Z \rho_b \frac{H}{H_p} = Z \rho_b \sqrt{\frac{\text{St} + \delta}{\delta}}.$$  \hspace{1cm} (13)

Here $Z$ is the ratio of the particle column density to the gas column density. In the limit of large particle sizes the mid-plane layer thickness is well approximated by $H_p = c_p/\Omega$, where $c_p$ is the random particle motion. This expression is obtained by associating the collision speed of large particles in equation (11) with their random speed (with $\alpha = \delta$) and inserting this into equation (12) in the limit $\text{St} \gg \delta$. The mass flux density of particles can then be written as

$$\mathcal{F} = c_p \rho_p = Z \rho_b H \Omega,$$  \hspace{1cm} (14)

independent of the collision speed as well as the degree of sedimentation, since the increased mid-plane density of larger particles is cancelled by the decrease in collision speeds. Hence the transition from Stokes numbers above unity to planetesimal sizes with significant gravitational cross sections is characterised by high collision speeds and slow, ordered growth ($\mathcal{F}$ is independent of size when $\mathcal{F}$ is constant) which does not benefit from increased particle sizes nor from decreased turbulent diffusion.

Mid-plane solids-to-gas ratios above unity are reached for $\text{St} = 1$ when $\delta < 10^{-4}$. Weaker stirring allows successively smaller particles to reach unity solids-to-gas ratio in the mid-plane. This marks an important transition to where particles exert a significant drag on the gas and become concentrated by streaming instabilities (Youdin and Goodman, 2005; Johansen et al., 2009a; Bai and Stone, 2010a,c). This effect will be discussed further in the next section.

4. PARTICLE CONCENTRATION

High local particle densities can lead to the formation of planetesimals by gravitational instability in the sedimented mid-plane layer. Particle densities above the Roche density

$$\rho_R = \frac{9 \Omega^2}{4\pi G} = 4.3 \times 10^{-7} \text{ g cm}^{-3} \left(\frac{r}{\text{AU}}\right)^{-3} = 315 \rho_b \left(\frac{r}{\text{AU}}\right)^{-1/4}$$  \hspace{1cm} (15)

are bound against the tidal force from the central star and can contract towards solid densities. The scaling with gas density assumes a distribution of gas according to the minimum mass solar nebula. Sedimentation increases the particle density in the mid-plane and can trigger bulk gravitational instabilities in the mid-plane layer (Goldreich and Ward, 1973). Sedimentation is nevertheless counteracted by turbulent diffusion (see section 3.4 and equation 12) and particle densities are prevented from reaching the Roche density by global turbulence or by mid-plane turbulence induced by the friction of the particle on the gas (Weidenschilling, 1980; Johansen et al., 2009a; Bai and Stone, 2010a).

High enough densities for gravitational collapse can nevertheless be reached when particles concentrate in the gas turbulence to reach the Roche density in local regions.

4.1. Passive concentration

The turbulent motion of gas in protoplanetary discs can lead to trapping of solid particles in the flow. The size of an optimally trapped particle depends on the length-scale and turn-over time-scale of the turbulent eddies. We describe here particle trapping progressively from the smallest scales unaffected by disc rotation (eddies) to the largest scales in near-perfect geostrophic balance between the pressure gradient and Coriolis accelerations (vortices and pressure bumps).

We consider rotating turbulent structures with length scale $\ell$, rotation speed $v_c$, and turn-over time $\tau_c = \ell/v_c$. These quantities are approximate and all factors of order unity are ignored in the following. Dust is considered to be passive particles accelerating towards the gas velocity in the friction time $\tau_1$, independent of the relative velocity. Particle trapping by streaming instabilities, where the particles play an active role in the concentration, is described in section 4.2. The three main mechanisms for concentrating particles in the turbulent gas flow in protoplanetary discs are sketched in Fig. 6.
4.1.1. Isotropic turbulence

On the smallest scales of the gas flow, where the Coriolis force is negligible over the turn-over time-scale of the eddies, the equation governing the structure of a rotating eddy is

\[ \frac{dv_e}{dt} = - \frac{1}{\rho} \frac{\partial P}{\partial r} \equiv f_P. \]  

(16)

Here \( f_P \) is the gas acceleration caused by the radial pressure gradient of the eddy. We use \( r \) as the radial coordinate in a frame centred on the eddy. The pressure must rise outwards, \( \partial P/\partial r > 0 \), to work as a centripetal force. In such low-pressure eddies the rotation speed is set by

\[ f_P = - \frac{v_e^2}{\ell}. \]  

(17)

Very small particles with \( \tau_f \ll t_e \) reach their terminal velocity

\[ v_p = - \tau_f f_P \]  

(18)
on a time-scale much shorter than the eddy turn-over time-scale. This gives

\[ v_p = - \tau_f f_P = \frac{\tau_f v_e^2}{\ell} = \frac{\tau_f}{t_e} v_e. \]  

(19)

The largest particles to reach their terminal velocity in the eddy turn-over time-scale have \( \tau_f \sim t_e \). This is the optimal particle size to be expelled from small-scale eddies and cluster in regions of high pressure between the eddies. Larger particles do not reach their terminal velocity before the eddy structure breaks down and reforms with a new phase, and thus their concentration is weaker.

Numerical simulations and laboratory experiments have shown that particles coupling at the turn-over time-scale of eddies at the Kolmogorov scale of isotropic turbulence experience the strongest concentrations (Squires and Eaton, 1991; Fessler et al., 1994). In an astrophysics context, such turbulent concentration of sub-mm-sized particles between small-scale eddies has been put forward to explain the narrow size ranges of chondrules found in primitive meteorites (Cuzzi et al., 2001), as well as the formation of asteroids by gravitational contraction of rare, extreme concentration events of such particles (Cuzzi et al., 2008). This model was nevertheless criticised by Pan et al. (2011) who found that efficiently concentrated particles have a narrow size range and that concentration of masses sufficiently large to form the primordial population of asteroids is unlikely.

4.1.2. Turbulence in rigid rotation

On larger scales of protoplanetary discs, gas and particle motion is dominated by Coriolis forces and shear. We first expand our particle-trapping framework to flows dominated by Coriolis forces and then generalise the expression to include shear.

In a gas rotating rigidly at a frequency \( \Omega \), the equilibrium of the eddies is now given by

\[ 2 \Omega v_e \frac{1}{\rho} \frac{\partial P}{\partial r} = - \frac{v_e^2}{\ell}. \]  

(20)

For slowly rotating eddies with \( v_e/\ell \ll \Omega \) we can ignore the centripetal term and get

\[ v_e = - \frac{f_P}{2\Omega}. \]  

(21)

High pressure regions have \( v_e < 0 \) (clockwise rotation), while low pressure regions have \( v_e > 0 \) (counter-clockwise rotation).

The terminal velocity of inertial particles can be found
The optimally trapped particle has a friction time that is proportional to the eddy turn-over time-scale, and thus all eddies trap particles of $\tau_\ell \approx 1/(2\Omega)$ most efficiently.

4.1.3. Turbulence with rotation and shear

Including both Keplerian shear and rotation the terminal velocity changes only slightly compared to equation (24),

$$v_p = \frac{2v_e}{(2\Omega_\tau_f)^{-1} + 2\Omega_\tau_f}. \quad (25)$$

Equilibrium structures in a rotating and shearing frame are axisymmetric pressure bumps surrounded by super-Keplerian/sub-Keplerian zonal flows (Fig. 6). The optimally trapped particle now has friction time $\Omega_\tau_f = 1$. This is in stark contrast to isotropic turbulence where each scale of the turbulence traps a small range of particle sizes with friction times similar to the turn-over time-scale of the eddy. The eddy speed is related to the pressure gradient through

$$v_e = -\frac{1}{2\Omega} \frac{1}{\rho} \frac{\partial P}{\partial r}. \quad (26)$$

Associating $v_e$ with $-\Delta v$ defined in equation (8), we recover the radial drift speed

$$v_r = -\frac{2\Delta v}{St^{-1} + St}. \quad (27)$$

However, this expression now has the meaning that particles drift radially proportional to the local value of $\Delta v$. Particles pile up where $\Delta v$ vanishes, i.e. in a pressure bump.

4.1.4. Origin of pressure bumps

Trapping of St$\sim$1 particles (corresponding to cm-sized pebbles to m-sized rocks and boulders, depending on the orbital distance, see Fig. 3) in pressure bumps in protoplanetary discs has been put forward as a possible way to cross the meter-barrier of planetesimal formation (Whipple, 1972; Haghighipour and Boss, 2003). Rice et al. (2004) identified this mechanism as the cause of particle concentration in spiral arms in simulations of self-gravitating protoplanetary discs. Shearing box simulations of turbulence caused by the magnetorotational instability (Balbus and Hawley, 1991) show the emergence of long-lived pressure bumps surrounded by a super-Keplerian/sub-Keplerian zonal flow (Johansen et al., 2009b; Simon et al., 2012). Similarly strong pressure bumps have been observed in global simulations (Fromang and Nelson, 2005; Lyra et al., 2008b).

High-pressure anticyclonic vortices concentrate particles in the same way as pressure bumps (Barge and Sommeria, 1995). Vortices may arise naturally through the baroclinic instability driving in the global entropy gradient (Klahr and Bodenheimer, 2003), although the expression of the baroclinic instability for long cooling times and in the presence of other sources of turbulence is not yet clear (Lesur and Papaloizou, 2010; Lyra and Klahr, 2011; Raettig et al., 2013).

Pressure bumps can also be excited by a sudden jump in the turbulent viscosity or by the tidal force of an embedded planet or star. Lyra et al. (2008a) showed that the inner and outer edges of the dead zone, where the ionisation degree is too low for coupling the gas and the magnetic field, extending broadly from 0.5 to 30 AU in nominal disc models (Dzyurkevich et al., 2013, chapter by Turner et al.), develop steep pressure gradients as the gas piles up in the low-viscosity region. The inner edge is associated with a pressure maximum and can directly trap particles. Across the outer edge of the dead zone the pressure transitions from high to low, hence there is no local maximum which can trap particles. The hydrostatic equilibrium nevertheless breaks into large-scale particle-trapping vortices through the Rossby wave instability (Li et al., 2001) in the variable-$\alpha$ simulations of Lyra et al. (2008a). Dzyurkevich et al. (2010) identified the inner pressure bump in resistive simulations of turbulence driven by the magnetorotational instability. The jump in particle density, and hence in ionisation degree, at the water snow line has been proposed to cause a jump in the surface density and act as a particle trap (Kretke and Lin, 2007; Brauer et al., 2008b; Drgžkowska et al., 2013). However, the snow line pressure bump remains to be validated in magnetohydrodynamical simulations.

Lyra et al. (2009) found that the edge of the gap carved by a Jupiter-mass planet develops a pressure bump which undergoes Rossby wave instability. The vortices concentrate particles very efficiently. Observational evidence for a large-scale vortex structure overdense in mm-sized pebbles was found by van der Marel et al. (2013). These particles have approximately Stokes number unity at the radial distance of the vortex, in the transitional disc IRS 48. This discovery marks the first confirmation that dust traps exist in nature. The approximate locations of the dust-trapping mechanisms which have been identified in the literature are shown in Fig. 7.
4.2. Streaming instability

The above considerations of passive concentration of particles in pressure bumps ignore the back-reaction friction force exerted by the particles onto the gas. The radial drift of particles leads to outwards motion of the gas in the mid-plane, because of the azimuthal frictional pull of the particles on the gas.

Youdin and Goodman (2005) showed that the equilibrium streaming motion of gas and particles is linearly unstable to small perturbations, a result also seen in the simplified mid-plane layer model of Goodman and Pindor (2000). The eight dynamical equations (six for the gas and particle velocity fields and two for the density fields) yield eight linear modes, one of which is unstable and grows exponentially with time. The growth rate depends on both the friction time and the particle mass-loading. Generally the growth rate increases proportional to the friction time (up to \( St \sim 1 \)), as the particles enjoy increasing freedom to move relative to the gas. The dependence on the particle mass-loading is more complicated; below a mass-loading of unity the growth rate increases slowly (much more slowly than linearly) with mass-loading, but after reaching unity in dust-to-gas ratio the growth rate jumps by one or more orders of magnitude. The \( e \)-folding time-scale of the unstable mode is as low as a few orbits in the regime of high mass-loading.

The linear mode of the streaming instability is an exact solution to the coupled equations of motion of gas and particles, valid for very small amplitudes. This property can be exploited to test numerical algorithm for solving the full non-linear equations. Youdin and Johansen (2007) tested their numerical algorithm for two-way drag forces against two modes of the streaming instability, the modes having different wavenumbers and friction times. These modes have subsequently been used in other papers (e.g. Bai and Stone, 2010b) to test the robustness and convergence of numerical algorithms for the coupled dynamics of gas and solid particles.

The non-linear evolution of the streaming instability can be studied either with or without particle stratification. The case without particle stratification is closest to the linear stability analysis. In this case the mean particle mass-loading in the simulation domain must be specified, as well as the friction time of the particles. The initial gas and particle velocities are set according to drag force equilibrium (Nakagawa et al., 1986), with particle drifting in towards the star and gas drifting out.

In simulations not including the component of the stellar gravity towards the mid-plane, i.e. non-stratified simulations, high particle densities are reached mainly for particles with \( St = 1 \) (Johansen and Youdin, 2007; Bai and Stone, 2010b). At a particle mass-loading of 3 (times the mean gas density), the particle density reaches almost 1000 times the gas density. The overdense regions appear nearly axisymmetric in both local and semi-global simulations (Kowalik et al., 2013). Smaller particles with \( St = 0.1 \) reach only between 20 and 60 times the gas density when the particle mass-loading is unity of higher. Little or no concentration is found at a dust-to-gas ratio of 0.2 for \( St = 0.1 \) particles. Bai and Stone (2010b) presented convergence tests of non-stratified simulations in 2-D and found convergence in the particle density distribution functions at 1024^2 grid cells.
Fig. 8.— Particle concentration by the streaming instability is mainly elongated along the Keplerian flow direction, as seen in the left panel showing the column density of particles with \( \text{St} = 0.3 \) in a local frame where the horizontal axis represents the radial direction and the vertical axis the orbital direction. The substructure of the overdense region appears fractal with filamentary structure on many scales. The right panel shows the measured maximum particle density as a function of time, for three different grid resolutions. The metallicity is gradually increased from the initial \( Z = 0.01 \) to \( Z = 0.02 \) between 30 and 40 orbits, triggering strong particle concentration. Higher resolution resolves smaller embedded substructures and hence higher densities.

4.2.1. Stratified simulations

While non-stratified simulations are excellent for testing the robustness of a numerical algorithm and comparing to results obtained with different codes, stratified simulations including the component of the stellar gravity towards the mid-plane are necessary to explore the role of the streaming instability for planetesimal formation. In the stratified case the mid-plane particle mass-loading is no longer a parameter that can be set by hand. Rather its value is determined self-consistently in a competition between sedimentation and turbulent diffusion. The global metallicity \( Z = \Sigma_p / \Sigma_g \) is now a free parameter and the mass-loading in the mid-plane depends on the scale height of the layer and thus on the degree of turbulent stirring.

Stratified simulations of the streaming instability display a binary behaviour determined by the heavy element abundance. At \( Z \) around the solar value or lower, the mid-plane layer is puffed up by strong turbulence and shows little particle concentration. Bai and Stone (2010a) showed that the particle mid-plane layer is stable to Kelvin-Helmholtz instabilities thriving in the vertical shear of the gas velocity, so the mid-plane turbulence is likely a manifestation of the streaming instability which is not associated with particle concentration. Above solar metallicity very strong particle concentrations occur in thin and dominantly axisymmetric filaments (Johansen et al., 2009a; Bai and Stone, 2010a). The metallicity threshold for triggering strong clumping may be related to reaching unity particle mass-loading in the mid-plane, necessary for particle pile-ups by the streaming instability. The particle density can reach several thousand times the local gas density even for relatively small particles of \( \text{St} = 0.3 \) (Johansen et al., 2012a). Measurements of the maximum particle density as a function of time are shown in Fig. 8, together with a column density plot of the overdense filaments.

Particles of \( \text{St} = 0.3 \) reach only modest concentration in non-stratified simulations. The explanation for the higher concentration seen in stratified simulations may be that the mid-plane layer is very thin, on the order of 1\% of a gas scale height, so that slowly drifting clumps are more likely to merge in the stratified simulations. The maximum particle concentration increases when the resolution is increased, by a factor approximately four each time the number of grid cells is doubled in each direction. This scaling may arise from thin filamentary structures which resolve into thinner and thinner filaments at higher resolution. The overall statistical properties of the turbulence nevertheless remain unchanged as the resolution increases (Johansen et al., 2012a).

The ability of the streaming instability to concentrate particles depends on both the particle size and the local metallicity in the disc. Carrera, Johansen and Davies (in preparation) show that particles down to \( \text{St} = 0.01 \) are concentrated at a metallicity slightly above the solar value, while smaller particles down to \( \text{St} = 0.001 \) require significantly increased metallicity, e.g. by photoevaporation of gas or pile up of particles from the outer disc. Whether the streaming instability can explain the presence of mm-sized chondrules in primitive meteorites is still not known. Alternative models based on small-scale particle concentration
(Cuzzi et al., 2008) and particle sedimentation (Youdin and Shu, 2002; Chiang and Youdin, 2010; Lee et al., 2010) may be necessary.

4.3. Gravitational collapse

Planetesimals forming by gravitational contraction and collapse of the overdense filaments are generally found to be massive, corresponding to contracted radii between 100 km and 1000 km, depending on the adopted disc model (Johansen et al., 2007, 2009a, 2011, 2012a; Kato et al., 2012). Fig. 9 shows the results of a high-resolution simulation of planetesimal formation through streaming instabilities, with planetesimals from 50 to 200 km radius forming when the model is applied to $r = 3$ AU. The formation of large planetesimals was predicted already by Youdin and Goodman (2005) based on the available mass in the linear modes. The planetesimals which form increase in size with increasing disc mass (Johansen et al., 2012a), with super-Ceres-sized planetesimals arising in massive discs approaching Toomre instability in the gas (Johansen et al., 2011). Increasing the numerical resolution maintains the size the largest planetesimals, but allows smaller-mass clumps to condense out of the turbulent flow alongside their massive counterparts (Johansen et al., 2012a). Although most of the particle mass enters planetesimals with characteristic sizes above 100 km, it is an open question, which can only be answered with very-high-resolution computer simulations, whether there is a minimum size to the smallest planetesimals which can form in particle concentration models.

Strictly speaking simulations of particle concentration and self-gravity only find the mass spectrum of gravitationally bound clumps. Whether these clumps will contract to form one or more planetesimals is currently not known in details. Nesvorný et al. (2010) took bound particle clumps similar to those arising in numerical simulations and evolved in a separate $N$-body code with collisions and merging. The rotating clumps typically collapse to a binary planetesimal, with orbital properties in good agreement with the observed high binary fraction in the classical cold Kuiper belt (Noll et al., 2008). The limited mass resolution which is allowed in direct $N$-body simulations nevertheless highlights the necessity for further studies of the fate of collapsing pebble clumps. Information from laboratory experiments on coagulation, bouncing and fragmentation of particles is directly applicable to self-gravitating particle clumps as well (Güttler et al., 2010), and including realistic particle interaction in collapse simulations should be a high priority for future studies.

5. PARTICLE GROWTH

The size distribution of solid particles in protoplanetary discs evolves due to processes of coagulation, fragmentation, sublimation and condensation. The particle concentration mechanisms discussed in the previous section are only relevant for particle sizes from millimeters to meters, much larger than the canonical sub-micron-sized dust and ice grains which enter the protoplanetary disc. Hence growth to macroscopic sizes must happen in an environment of relatively uniform particle densities.

Primitive meteorites show evidence of thermal processing of protoplanetary disc solids, in the form of CAIs which likely condensed directly from the cooling gas near the proto-Sun and chondrules which formed after rapid heating and melting of pre-existing dust aggregates (section 2.2). Thermal processing must also have taken place near the water snow line where the continued process of sublimation and condensation can lead to the formation of large hail-like ice particles (Ros and Johansen, 2013); or more distantly around the CO snow line identified observationally at 30 AU around the young star TW Hya (Qi et al., 2013).

Direct sticking is nevertheless the most general mechanism for dust growth, as the densities in protoplanetary discs are high enough for such coagulation to be important over the life-time of the gaseous disc. Physically, the collisional evolution is determined by the following factors:

- The spatial density of particles. This will depend on both the location in the disc and on the degree of par-
particle concentration (see section 4).

- The collision cross section among particles, $\sigma_{\text{col}}$. An appropriate assumption is to take spherical particles, for which $\sigma_{\text{col},ij} = \pi(R_i + R_j)^2$ where $R_i$ and $R_j$ are simply the radii of two particles in the absence of gravitational focusing.
- The relative velocity among the particles, $\delta v_{ij}$. The relative velocity affects both the collision rate and the collision outcome.
- The collision outcome, which is determined by the collision energy, porosity, and other collision parameters like the impact parameter.

The product of $\delta v_{ij}$ and $\sigma_{ij}$, which together with the particle density determine the collision rate, is most often referred to as the collision kernel $K_{ij}$.

5.1. Numerical approaches the coagulation equation

The evolution of the dust size distribution is most commonly described by the Smoluchowski equation (Smoluchowski, 1916), which reads:

$$\frac{dn(m)}{dt} = \frac{1}{2} \int \int K(m',m-n')n(m')n(m-n')n(m) - n(m) \int \int K(m',m')n(m')$$

where $n(m)$ is the particle number density distribution (i.e., $n(m)\,dm$ gives the number density of particles between mass $m$ and $m + dm$). The terms on the right-hand side of equation (28) simply account for the particles of mass $m$ that are removed due to collisions with any other particle and those gained by a collision between two particles of mass $m'$ and $m - m'$ (the factor of 1/2 prevents double counting). In numerical applications one most often subdivides the mass grid in bins (often logarithmically spaced) and then counts the interactions among these bins.

However, equation (28), an already complex integro-differential equation, ignores many aspects of the coagulation process. For example, collisional fragmentation is not included and the dust spatial distribution is assumed to be homogeneous (that is, there is no local concentration). Equation (28) furthermore assumes that each particle pair $(m, m')$ collides at a single velocity and ignores the effects of a velocity distribution. A new set of collision rates is computed, after which the procedure repeats itself.

Since there are many more dust grains in a disc than any computer can handle, the computational particles should be chosen such to accurately represent the physical size distribution. A natural choice is to sample the mass of the distribution (Zsom and Dullemond, 2008). This method also allows fragmentation to be naturally incorporated. The drawback however is that the tails of the size distribution are not well resolved. The limited dynamic range is one of the key drawbacks of MC-methods. (2008) have described a solution, but at the expense of a more complex algorithm.

5.2. Laboratory experiments

Numerical solutions to the coagulation equation rely crucially on input from laboratory experiments on the outcome of collisions between dust particles. Collision experiments in the laboratory as well as under microgravity conditions over the past 20 years have proven invaluable for the modeling of the dust evolution in protoplanetary discs. Here we briefly review the state-of-the-art of these experiments (more details are given in the chapter by Testi et al.). A detailed physical model for the collision behavior of silicate aggregates of all masses, mass ratios and porosities can be found in G"uttler et al. (2010) and recent modifications are published by Kothe et al. (2013).

The three basic types of collisional interactions between dust aggregates are:

1. Direct collisional sticking: when two dust aggregates collide gently enough, contact forces (e.g., van der Waals forces) are sufficiently strong to bind the aggregates together. Thus, a more massive dust aggregate has formed. Laboratory experiments summarized by Kothe et al. (2013) suggest that there is
a mass-velocity threshold \( v \propto m^{-3/4} \) for spherical medium-porosity dust aggregates such that only dust aggregates less massive and slower than this threshold can stick. However, hierarchical dust aggregates still stick at higher velocities/masses than given by the threshold for homogeneous dust aggregates (Kothe et al., 2013).

2. Fragmentation: when the collision energy is sufficiently high, similar-sized dust aggregates break up so that the maximum remaining mass decreases under the initial masses of the aggregates and a power-law type tail of smaller-mass fragments is produced. The typical fragmentation velocity for dust aggregates consisting of \( \mu \text{m} \)-sized silicate monomer grains is \( 1 \text{ m s}^{-1} \).

3. Bouncing: above the sticking threshold and below the fragmentation limit, dust aggregates bounce off one another. Although the conditions under which bouncing occurs are still under debate (see sections 6 and 7), some aspects of bouncing are clear. Bouncing does not directly lead to further mass gain so that the growth of dust aggregates is stopped (“bouncing barrier”). Furthermore, bouncing leads to a steady compression of the dust aggregates so that their porosity decreases to values of typically 60% (Weidling et al., 2009).

Although the details of the three regimes are complex – regime boundaries are not sharp, and other more subtle processes exist – this simplified physical picture shows that in protoplanetary discs, in which the mean collision velocity increases with increasing dust-aggregate size (see Fig. 4), a maximum aggregate mass exists. Detailed numerical simulations using the dust-aggregate collision model by Gütler et al. (2010) find such a maximum mass (Zsom et al., 2010; Windmark et al., 2012b). For a minimum mass solar nebula model at 1 AU, the maximum dust-aggregate size is in the range of millimeters to centimeters (e.g. Zsom et al., 2010).

Even if the bouncing barrier can be overcome, another perhaps even more formidable obstacle will present itself at the size scale where collision velocities reach \(~1 \text{ m/s}\) (Fig. 4). At this size collisions among two silicate, similar-size, dust particles are seen to fragment, rather than accrete. Consequently, growth will stall and the size distribution will, like with the bouncing case, settle into a steady-state (Birnstiel et al., 2010, 2011, 2012). Highly porous icy aggregates may nevertheless still stick at high collision speeds (see section 7).

6. GROWTH BY MASS TRANSFER

Over the last two decades a large effort has been invested to study experimentally the collisions among dust particles and to assess the role of collisions in planetesimal formation. As was seen in section 5 a plethora of outcomes – sticking, compaction, bouncing, fragmentation – is possible depending mainly on the collision velocity and the particle size ratio. Using the most updated laboratory knowledge of collision outcomes, Zsom et al. (2010, 2011) solved the coagulation equation with a Monte Carlo method and found that in the ice-free, inner disc regions the dust size distribution settles into a state dominated by mm-size particles. Further growth is impeded because collisions among two such mm-size particles will mainly lead to bouncing and compaction.

The real problem behind this “bouncing barrier” may, counter-intuitively, not be the bouncing but rather the absence of erosion and fragmentation events. Based upon laboratory experiments (Wurm et al., 2005; Teiser and Wurm, 2009b,a; Kothe et al., 2010; Teiser et al., 2011; Meisner et al., 2013), it has become clear that above the fragmentation limit, the collision between two dust aggregates can lead to a mass gain of the larger (target) aggregate if the smaller (projectile) aggregate is below a certain threshold. For impact velocities in the right range and relatively small projectile aggregates, up to 50% of the mass of the projectile can be firmly transferred to the target. This process has been shown to continue after multiple collisions with the target and under a wide range of impact angles (see Fig. 10).

Growth by mass transfer requires a high fraction of the dust mass present in small dust aggregates which can be swept up by the larger particles. The number of small dust
aggregates can be maintained by catastrophic collisions between larger bodies. A simple two-component model for such coagulation-fragmentation growth was developed by Johansen et al. (2008). They divided the size distribution into small particles and large particles and assumed that (1) a small particle will stick to a large particle, (2) a small particle will bounce off another small particle, and (3) a large particle will shatter another large particle. Small particles collide with large particles at speed \( v_{12} \), while large particles collide with each other at speed \( v_{22} \). Under these conditions an equilibrium in the surface density ratio of large and small particles, \( \Sigma_1/\Sigma_2 \), can be reached, with equal mass flux from large to small particles by fragmentation as from small to large particles by mass transfer in collisions.

In the equilibrium state where the mass flux from large to small particles balances the mass flux from small to large particles, the growth rate of the large particles is

\[
\dot{R} = \frac{\Sigma_p/(\sqrt{2\pi H_1})}{\rho_\bullet} \frac{v_{22}}{[2/(1 + H_1^2/H_2^2)]^{1/2} + 4v_{22}/v_{12}}. \tag{29}
\]

Here \( H_1 \) and \( H_2 \) are the scale heights of the small particles and the large particles, respectively, and \( \Sigma_p = \Sigma_1 + \Sigma_2 \) is the total column density of the dust particles. The complexity of the equation arises from the assumption that the small grains are continuously created in collisions between the large particles. Setting \( H_1 = H \) and \( H_2 \ll H \) the equations are valid for a very general sweep-up problem where a few large particles sweep up a static population of small particles, resulting in growth of the large particles at rate

\[
\dot{R} = \frac{\epsilon Z \rho_\bullet}{4\rho_\bullet} v_{12} = 2.7 \text{ mm yr}^{-1} \left( \frac{\epsilon}{0.5} \right) \left( \frac{r}{\text{AU}} \right)^{-2.75} \times \left( \frac{Z}{0.01} \right) \left( \frac{\rho_\bullet}{1 \text{ g cm}^{-3}} \right)^{-1} \left( \frac{v_{12}}{50 \text{ m s}^{-1}} \right). \tag{30}
\]

We introduced here a sticking coefficient \( \epsilon \) which measures either the sticking probability or the fraction of mass which is transferred from the projectile to the target. The growth rate is constant and depends only on the particle size through \( v_{12} \). If the small particles are smaller than Stokes number unity and the large particles are larger, then we can assume that the flux of small particles is carried with the sub-Keplerian wind, \( v_{12} \approx \Delta v \), and solve equation (30) for the time-scale to cross the radial drift barrier. The end of the radial drift barrier can be set at reaching roughly \( St = 10 \). This Stokes number corresponds in the Epstein regime to the particle size

\[
R_{10}^{(Ep)} = \frac{10H \rho_\bullet}{\rho_\bullet}. \tag{31}
\]

In the Stokes regime the particle size is

\[
R_{10}^{(St)} = \sqrt{\frac{90H \rho_\bullet \lambda}{4\rho_\bullet}}. \tag{32}
\]

Fig. 3 shows that compact ice particles grow to Stokes number 10 in the Epstein regime outside of \( r = 5 \) AU and in the Stokes or non-linear regimes inside of this radius. Fluffy ice particles with a very low internal density of \( \rho_\bullet = 10^{-6} \text{ g cm}^{-3} \) grow to Stokes number 10 in the non-linear or quadratic regimes in the entire extent of the disc. In the Epstein regime the time to grow to Stokes number 10, \( \tau_{10} = R_{10}/\dot{R} \), is independent of both material density and gas density.

\[
\tau_{10}^{(Ep)} = \frac{40H\rho_\bullet}{\epsilon Z v_{12}} \approx 25000 \text{ yr} \left( \frac{\epsilon}{0.5} \right)^{-1} \left( \frac{r}{\text{AU}} \right)^{5/4}. \tag{33}
\]

The growth-time is much shorter in the Stokes regime where the Stokes number is proportional to the squared particle radius, giving

\[
\tau_{10}^{(St)} = \frac{\sqrt{360H \rho_\bullet (m/\sigma)}}{\epsilon Z \rho_\bullet v_{12}} \approx 500 \text{ yr} \times \left( \frac{\epsilon}{0.5} \right)^{-1} \left( \frac{\rho_\bullet}{1 \text{ g cm}^{-3}} \right)^{1/2} \left( \frac{r}{\text{AU}} \right)^{27/8}. \tag{34}
\]

with a direct dependence on both material density and gas density. Here \( m \) and \( \sigma \) are the mass and collisional cross section of a hydrogen molecule. Regarding the sticking coefficient, Wurm et al. (2005) find mass transfer efficiencies of \( \epsilon \approx 0.5 \) for collision speeds up to \( v_{12} = 25 \) m/s, hence we have used \( \epsilon = 0.5 \) as a reference value in the above equations. It is clearly advantageous to have a low internal density and cross the radial drift barrier in the Stokes regime (Okuzumi et al., 2012). The non-linear and quadratic regimes are similarly beneficial for the growth, with the time-scale for crossing the meter barrier in the quadratic drag force regime approximately a factor \( 6c_\bullet/\Delta v \) faster than in the Epstein regime (equation 33).

The two-component model presented in Johansen et al. (2008) is analytically solvable, but very simplified in the collision physics. Windmark et al. (2012b) presented solutions to the full coagulation equation including a velocity-dependent mass transfer rate. They showed that artificially injected seeds of centimeter sizes can grow by sweeping up smaller mm-sized particles which are stuck at the bouncing barrier (Zsom et al., 2010), achieving at 3 AU growth to 3 meters in 100,000 years and to 100 meters in 1 million years. The time-scale to grow across the radial drift barrier is broadly in agreement with the simplified model presented above in equation (34). This time is much longer than the radial drift time-scale which may be as short as 100–1,000 years. It is therefore necessary to invoke pressure bumps to protect the particles from radial drift while they grow slowly by mass transfer. However, in that case the azimuthal drift of small particles vanishes and a relatively high turbulent viscosity of \( \alpha = 10^{-2} \) must be used to regain the flux onto the large particles. The pressure bumps must additionally be very long-lived, on the order of the life-time of the protoplanetary disc.

While the discovery of a pathway for growth by mass transfer at high-speed collisions is a major experimental
breakthrough, it seems that this effect can not in itself lead to widespread planetesimal formation, unless perhaps in the innermost part of the protoplanetary disc where dynamical time-scales are short (equation 34). The planetesimal sizes which are reached within a million years are also quite small, on the order of 100 meters. These sizes are too small to undergo gravitational focusing even in weakly turbulent discs (see Fig. 5). An additional concern is the erosion of the preplanetesimal by tiny dust grains carried with the sub-Keplerian wind (Schräpler and Blum, 2011; Seizinger et al., 2013).

Particles stuck at the bouncing barrier are at the lower end of the sizes that can undergo particle concentration and gravitational collapse. However, if a subset of “lucky” particles experience only low-speed, sticking collisions and manage to grow past the bouncing barrier, this leads eventually to the formation of a population of cm-m-sized particles that could concentrate in pressure bumps or participate in the streaming instability (Windmark et al., 2012a; Garaud et al., 2013).

7. FLUFFY GROWTH

The last years have seen major improvements in N-body molecular-dynamics simulations of dust aggregate collisions by a number of groups (e.g. Wada et al., 2008, 2009, 2011; Paszun and Dominik, 2008, 2009; Seizinger and Kley, 2013). These N-body simulations show that fluffy aggregates have the potential to overcome barriers in dust growth. In N-body simulations of dust aggregates, all surface interactions between monomers in contact in the aggregates are calculated, by using a particle interaction model.

N-body simulations of dust aggregates have some merits compared to laboratory experiments. Precise information such as the channels through which the collisional energy is dissipated can be readily obtained and subsequent data analysis is straightforward. Another advantage of N-body simulations is that one can study highly fluffy aggregates, which are otherwise crushed under the Earth’s gravity. On the other hand, an accurate interaction model of constituent particles is required in N-body simulations of dust aggregates. In the interaction model used in most N-body simulations of dust aggregates, the constituent particles are considered as adhesive elastic spheres. The adhesion force between them is described by the JKR theory (Johnson et al., 1971). As for tangential resistive forces against sliding, rolling, or twisting motions, Dominik and Tielens’s model is used (Dominik and Tielens, 1995, 1996, 1997; Wada et al., 2007). In order to reproduce the results of laboratory experiments better, the interaction model must be calibrated against laboratory experiments (Paszun and Dominik, 2008; Seizinger et al., 2012; Tanaka et al., 2012).

In protoplanetary discs, dust aggregates are expected to have fluffy structures with low bulk densities if impact compaction is negligible (e.g. Okuzumi et al., 2009, 2012; Zsom et al., 2010, 2011). Especially in the early stage of dust growth, low-speed impacts result in hit-and-stick of dust aggregates, with little or no compression. This growth mode makes fluffy aggregates with a low fractal dimension of ~2. Thus it is necessary to examine the outcome of collisions between fluffy dust aggregates and the resulting compression. The results of N-body simulations can then be used in numerical models of the coagulation equation including the evolution of the dust porosity.

7.1. Critical impact velocity for dust growth

Dominik and Tielens (1997) first carried out N-body simulation of aggregate collisions, using the particle interaction model they constructed. They also derived a simple recipe for outcomes of aggregate collisions from their numerical results, though they only examined head-on collisions of two-dimensional small aggregates containing as few as 40 particles. The collision outcomes are classified into hit-and-stick, sticking with compression, and catastrophic disruption, depending on the impact energy. Bouncing of two aggregates was not observed in their simulations.

According to the DT (Dominik and Tielens, 1997) recipe, growth is possible at collisions with \( E_{\text{imp}} < A n_k E_{\text{break}} \), where \( A \sim 10 \) and \( n_k \) is the total number of contacts in two colliding aggregates. For relatively fluffy aggregates, \( n_k \) is approximately equal to the total number of constituent particles.
particles in the two aggregates, $N$. The energy for breaking one contact between two identical particles, $E_{\text{break}}$, is given by (e.g. Chokshi et al., 1993; Wada et al., 2007)

$$E_{\text{break}} = 2\gamma r^4 \left(1 - \nu^2 / \varepsilon^2 \right)^{1/3},$$

(35)

where $r$, $\gamma$, $\varepsilon$, and $\nu$ are the radius, the surface energy, Young's modulus, and Poisson's ratio of constituent particles, respectively. For icy particles the breaking energy is obtained as $5.2 \times 10^{-10} (r / 0.1 \mu m)^{4/3}$ erg while the value for silicate particles is $1.4 \times 10^{-11} (r / 0.1 \mu m)^{4/3}$ erg, so icy aggregates are much more sticky than silicate aggregates.

Using the impact velocity, $v_{\text{imp}}$, the impact energy is expressed as $(Nm/4)v_{\text{imp}}^2/2$, where $m$ is the mass of a constituent particle and $Nm/4$ is the reduced mass of two equal-sized aggregates. Substituting this expression into the above energy criterion for dust growth and noting that $n_k \approx N$, we obtain the velocity criterion as

$$v_{\text{imp}} < \sqrt{8A E_{\text{break}} / m}.$$  

(36)

Here $A$ is a dimensionless parameter that needs to be calibrated with experiments. Note that this velocity criterion is dependent only on properties of constituent particles and, thus, independent of the mass of the aggregates.

To evaluate the critical impact velocity for growth more accurately, Wada et al. (2009) performed $N$-body simulations of large aggregates made of up to $\sim 10^4$ sub-micron icy particles, including off-set collisions (see Fig. 11). They consider two kinds of aggregate structure, the so-called CCA and PCA clusters. The CCA (Cluster-Cluster Agglomeration) clusters have an open structure with a fractional dimension of 2 while the PCA (Particle-Cluster Agglomeration) clusters have a fractal dimension of 3 and a volume filling factor of 0.15. The PCA clusters are rather compact, compared to the CCA. Since dust aggregates are expected to be much more compact than CCA clusters due to compression (see section 7.2), the growth and disruption process of PCA clusters is of particular importance for elucidating the planetesimal formation.

For PCA clusters (composed of 0.1 $\mu$m-sized icy particles), the critical impact velocity is obtained as 60 m/s from their simulations, independent of the aggregate mass within the mass range examined in the simulations of Wada et al. (2009). This indicates that icy dust aggregates can circumvent the fragmentation barrier and grow towards planetesimal sizes via collisional sticking. Note that the critical velocity actually increases with the aggregate mass when only head-on collisions are considered, as seen in Fig. 11. For an accurate evaluation of the critical velocity for growth, offset collisions should be considered as well, as in Wada et al. (2009).

The above result also fixes the corresponding constant $A$ in the critical energy at 30. By using $E_{\text{break}}$ of silicate particles in equation (36) with $A = 30$, the critical velocity for growth of silicate aggregates is obtained as $v_{\text{imp}} = 1.3 (r / 0.6 \mu m)^{-5/6}$ m/s. It agrees well with the laboratory experiments and numerical simulations for small silicate aggregates (e.g. Blum and Warm, 2000; Gütler et al., 2010; Paszun and Dominik, 2009; Seizinger and Kley, 2013).

In the case of CCA clusters, Wada et al. (2009) found that the constant $A$ is the same as in the DT recipe ($\approx 10$). Considering much smaller volume filling factors of CCA clusters than PCA, it indicates that the critical velocity is only weakly dependent on the volume filling factor.

7.2. Compression of dust aggregates

The DT recipe does not describe the amount of changes in the porosity (or the volume filling factor) at aggregate collisions. The first attempt to model porosity changes was done by Ormel et al. (2007), using simple prescriptions for the collision outcome of porosity.

Wada et al. (2008) examined compression at head-on collisions of equal-sized PCA clusters, using a high number of $N$-body simulations. Compression (or restructuring) of an aggregate occurs through rolling motions between constituent particles. Thus it is governed by the rolling energy $E_{\text{roll}}$ (i.e., the energy for rolling of a particle over a quarter of the circumference of another particle in contact). At low-energy impacts with $E_{\text{imp}} \gtrsim E_{\text{roll}}$, aggregates just stick to each other, as indicated by the DT recipe. For higher-energy impacts, the resultant aggregates are compressed, depending on the impact energy. The radius of resulting aggregates, $R$, is fitted well with the power-law function

$$R \approx 0.8 \left[ E_{\text{imp}} / (N E_{\text{roll}}) \right]^{-0.1} N^{1/2} r^5,$$

(37)

where $r$ is the radius of a monomer and $N$ the number of constituent particles. Analysing the structure of the compressed aggregates, it can be shown that the compressed aggregates have a fractal dimension of 2.5. This fractal dimension is consistent with equation (37) since equation (37) gives the relation of $N \propto R^{2.5}$ for maximally compressed aggregates (with $E_{\text{imp}} \approx N E_{\text{roll}}$). The compression observed in the $N$-body simulation is less much extreme than in the simple porosity model used by Ormel et al. (2007) because of the low fractal dimension of 2.5.

While Wada et al. (2008) examined the compression in a single collision, dust aggregates will be gradually compressed by successive collisions in realistic systems. In order to examine such a gradual compression process during growth, Suyama et al. (2008) performed $N$-body simulations of sequential collisions of aggregates. Even after multiple collisions, the compressed aggregates maintain a fractal dimension of 2.5. Suyama et al. (2012) further extend their porosity model to unequal-mass collisions.

Okuzumi et al. (2012) applied the porosity evolution model of Suyama et al. (2012) to dust growth outside of the snow line in a protoplanetary disc. As a first step, collisional fragmentation is neglected because a relatively high impact velocity ($\sim 60$ m/s) is required for significant disruption of icy dust aggregates. They found that dust particles evolve into highly porous aggregates (with bulk densities
much less than 0.1 g/cm³) even if collisional compression is taken into account. This is due to the ineffective compression at aggregate collisions in the porosity model of Suwa et al. (2012). Another important aspect of fluffy aggregates is that they cross the radial drift barrier in the Stokes or nonlinear drag force regimes (see Fig. 3 and discussion in section 6) where it is easier to grow to sizes where radial drift is unimportant. This mechanism accelerates dust growth at the radial drift barrier effectively and enables fluffy aggregates to overcome this barrier inside 10 AU (Okuzumi et al., 2012; Kataoka et al., 2013).

7.3. Bouncing condition in N-body simulations

Until 2011, bouncing events were not reported in N-body simulations of aggregate collisions, while bouncing events have been frequently observed in laboratory experiments (e.g. Blum and Münch, 1993; Langkowski et al., 2008; Weiding et al., 2009, 2012). Wada et al. (2011) found bouncing events in their N-body simulations of both icy and silicate cases for rather compact aggregates with filling factor $\phi \geq 0.35$. Such compact aggregates have a relatively large coordination number (i.e., the mean number of particles in contact with a particle), which inhibits the energy dissipation through the rolling deformation and helps bouncing. Seizinger and Kley (2013) further investigated the bouncing condition and proposed that the more realistic condition is $\phi > 0.5$.

This critical volume filling factor for bouncing is a few times as large as that in the laboratory experiments (e.g. Langkowski et al., 2008). The origin of this discrepancy between these two approaches is not clear yet. From a qualitative point of view, however, laboratory experiments show that fluffy silicate aggregates with $\phi < 0.1$ tend to stick to each other (Blum and Wurm, 2000; Langkowski et al., 2008; Kothe et al., 2013). This qualitative trend is consistent with the N-body simulations. In protoplanetary discs, dust aggregates are expected to be highly porous. In the growth of such fluffy aggregates, the bouncing barrier would not be a strong handicap.

8. TOWARDS A UNIFIED MODEL

Three main scenarios for the formation of planetesimals have emerged in the last years: (1) formation via coagulation-fragmentation cycles and mass transfer from small to large aggregates (Wurm et al., 2005; Windmark et al., 2012b), (2) growth of fluffy particles and subsequent compactification by self-gravity (Wada et al., 2008, 2009), and (3) concentration of pebbles in the turbulent gas and gravitational fragmentation of overdense filaments (Johansen et al., 2007, 2009a; Kato et al., 2012).

The collision speeds at which planetesimal-sized bodies form are typically 50 m/s in models (1) and (2), as small dust aggregates are carried onto the growing planetesimal with the sub-Keplerian wind, or even higher if the turbulent density fluctuations are strong. During the gravitational collapse of pebble clouds, formation mechanism (3) above, the collision speeds reach a maximum of the escape speed of the forming body. In the cometary formation zone any subsequent impacts by high-speed particles, brought in with the sub-Keplerian flow, will at most add a few meters of compact debris to the pebble pile over the life-time of the solar nebula (equation 30 applied to $r = 10$ AU).

This difference in impact velocity leads to substantial differences in the tensile strengths of the planetesimals/cometesimals. Due to the relatively high impact speeds in models (1) and (2), the growing bodies are compacted and possess tensile strengths on the order of 1–10 kPa (Blum et al., 2006). Skorov and Blum (2012) pointed out that these tensile-strength values are too high to explain the continuous dust emission of comets approaching the Sun and favor model (3) above, for which the tensile strengths of loosely packed mm-cm pebbles should be much smaller.

Skorov and Blum (2012) base their comet-nucleus model on a few assumptions: (a) dust aggregates in the inner parts of the protoplanetary disc can only grow to sizes of mm-cm; (b) turbulent diffusion can transport dust aggregates to the outer disc; (c) in the outer disc, dust and ice aggregates become intermixed with a dust-to-ice ratio found in comets; (d) cometesimals form via gravitational instability of overdense regions of mm-cm-sized particles which will form planetesimals with a wide spectrum of masses. While km-scale planetesimals have not yet been observed to form in hydrodynamical simulations of planetesimal formation, this may be an artifact of the limited numerical resolution (see discussion in section 4.3).

If the comet nuclei as we find them today have been dormant in the outer reaches of the Solar System since their formation and have not been subjected to intense bombardment and aqueous or thermal alteration, then they today represent the cometesimals of the formation era of the Solar System. Skorov and Blum (2012) derive for their model tensile strengths of the ice-free dusty surfaces of comet nuclei

$$T = T_0 \left( \frac{r}{1 \text{ mm}} \right)^{-2/3},$$

with $T_0 = 0.5$ Pa and $r$ denoting the radius of the dust and ice aggregates composing the cometary surface.

These extremely low tensile-strength values are, according to the thermophysical model of Skorov and Blum (2012), just sufficiently low to explain a continuous outgassing and dust emission of comet nuclei inside a critical distance to the Sun. This model provides strong indication that km-sized bodies in the outer Solar System were formed by the gravitational contraction of smaller sub-units (pebbles) at relatively low (order of m/s) velocities. Such pebble-pile planetesimals are in broad agreement with the presence of large quantities of relatively intact mm-sized chondrules and CAIs in chondrites from the asteroid belt.

The cometesimals could thus represent the smallest bodies which formed by gravitational instability of mm-cm-sized icy/rocky pebbles. These pebbles will continue to collide inside the gravitationally collapsing clump. Low-
mass clumps experience low collision speeds and contract as kinetic energy is dissipated in inelastic collisions, forming small planetesimals which are made primarily of pristine pebbles. Clumps of larger mass have higher collision speeds between the constituent pebbles, and hence growth by coagulation-fragmentation cycles inside massive clumps determines the further growth towards one or more solid bodies which may go on to differentiate by decay of short-lived radionuclides. This way coagulation is not only important for forming pebbles that can participate in particle concentration, but continues inside of the collapsing clumps to determine the birth size distribution of planetesimals.

There is thus good evidence that pebbles are the primary building blocks of planetesimals both inside the ice line (today’s asteroids) and outside the ice line (today’s comets and Kuiper belt objects). This highlights the need to understand pebble formation and dynamics better. The bouncing barrier is a useful way to maintain a high number of relatively small pebbles in the protoplanetary disc (Zsom et al., 2010). As pebbles form throughout the protoplanetary disc, radial drift starts typically when reaching mm-cm sizes. Such small pebbles drift relatively fast in the outer disc but slow down significantly (to speeds of order 10 cm/s) in the inner few AU. Pebbles with a high ice fraction fall apart at the ice line around 3 AU, releasing refractory grains, which can go on to form new (chondrule-like) pebbles inside the ice line (Sirono, 2011), as well as water vapor which boosts formation of large, icy pebbles outside the ice line (Ros and Johansen, 2013). Similar release of refractories and rapid pebble growth will occur at ice lines of more volatile species like CO (Qi et al., 2013). In the optically thin very inner parts of the protoplanetary disc, illumination by the central star leads to photophoresis which causes chondrules to migrate outwards (Loesche et al., 2013). Radial drift of pebbles will consequently not lead to a widespread depletion of planetesimal building blocks from the protoplanetary disc. Instead pebble formation can be thought of as a continuous cycle of formation, radial drift, destruction and reformation.

We propose therefore that particle growth to planetesimal sizes starts unaided by self-gravity, but after reaching Stokes numbers roughly between 0.01 and 1 proceeds inside self-gravitating clumps of pebbles (or extremely fluffy ice balls with similar aerodynamic stopping times). In this picture coagulation and self-gravity are not mutually exclusive alternatives but rather two absolutely necessary ingredients in the multifaceted planetesimal formation process.

Acknowledgments. We thank the referee for useful comments that helped improve the manuscript. AJ was supported by the Swedish Research Council (grant 2010-3710), the European Research Council under ERC Starting Grant agreement 278675-PEBBLE2PLANET and by the Knut and Alice Wallenberg Foundation. MB acknowledges funding from the Danish National Research Foundation (grant number DNRF97) and by the European Research Council under ERC Consolidator grant agreement 616027-STARDUST2ASTEROIDS. CWO acknowledges support for this work by NASA through Hubble Fellowship grant No. HST-HF-51294.01-A awarded by the Space Telescope Science Institute, which is operated by the Association of Universities for Research in Astronomy, Inc., for NASA, under contract NAS 5-26555. HR was supported by the Swedish National Space Board (grant 74/10:2) and the Polish National Science Center (grant 2011/01/B/ST9/05442). This work was granted access to the HPC resources of JuRoPA/FZJ, Stokes/ICHEC and RZG P6/RZG made available within the Distributed European Computing Initiative by the PRACE-2IP, receiving funding from the European Community’s Seventh Framework Programme (FP7/2007-2013) under grant agreement no. RI-283493.

REFERENCES
Botlife W. F. et al. (2005) Icarus, 175, 111.
D’ Alessio P. et al. (2005) in: Chondrites and the Protoplanetary Disk, vol. 341 of Astronomical Society of the Pacific Con-