Mon. Not. R. Astron. Soc. **000**, 1–42 (2011) Printed 20 May 2014 (MN LATEX style file v2.2)

Hybrid methods in planetesimal dynamics (I) : Description of a new composite algorithm

P. Glaschke[1] , P. Amaro-Seoane[⋆2,3,4] & R. Spurzem[3,1,5]
[1] *Astronomisches Rechen-Institut, Mönchhofstraße 12-14, Zentrum für Astronomie, Universität Heidelberg, Germany*
[2] *Max Planck Institut für Gravitationsphysik (Albert-Einstein-Institut), D-14476 Potsdam, Germany*
[3] *National Astronomical Observatories of China, Chinese Academy of Sciences, 20A Datun Lu, Chaoyang District, 100012, Beijing, China*
[4] *Institut de Ciències de l'Espai (CSIC-IEEC), Campus UAB, Torre C-5, parells, 2[na] planta, ES-08193, Bellaterra, Barcelona, Spain*
[5] *Kavli Institute for Astronomy and Astrophysics, Peking University, China*

arXiv:1105.6094v1 [astro-ph.EP] 30 May 2011

draft 20 May 2014

ABSTRACT

The formation and evolution of protoplanetary systems, the breeding grounds of planet formation, is a complex dynamical problem that involves many orders of magnitudes. To serve this purpose, we present a new hybrid algorithm that combines a Fokker-Planck approach with the advantages of a pure direct-summation $N-$body scheme, with a very accurate integration of close encounters for the orbital evolution of the larger bodies with a statistical model, envisaged to simulate the very large number of smaller planetesimals in the disc. Direct-summation techniques have been historically developped for the study of dense stellar systems such as open and globular clusters and, within some limits imposed by the number of stars, of galactic nuclei. The number of modifications to adapt direct-summation $N-$body techniques to planetary dynamics is not undemanding and requires modifications. These include the way close encounters are treated, as well as the selection process for the "neighbour radius" of the particles and the extended Hermite scheme, used for the very first time in this work, as well as the implementation of a central potential, drag forces and the adjustment of the regularisation treatment. For the statistical description of the planetesimal disc we employ a Fokker-Planck approach. We include dynamical friction, high- and low-speed encounters, the role of distant encounters as well as gas and collisional damping and then generalise the model to inhomogenous discs. We then describe the combination of the two techniques to address the whole problem of planetesimal dynamics in a realistic way via a transition mass to integrate the evolution of the particles according to their masses.

Key words: protoplanetary discs, planets and satellites: dynamical evolution and stability, methods: numerical, methods: N-body, methods: statistical

1 INTRODUCTION

The formation of a planetary system is closely related to the formation of the host star itself. Cool molecular clouds collapse and fragment into smaller substructures which are the seeds for subsequent star formation. Angular momentum conservation in the forming clumps forces the infalling matter into a disc-like structure. The subsequent viscous evolution of the disc leads to a transport of angular momentum which channels gas to the protostar in the centre. These protoplanetary discs are the birth place of planets (for a detailed review see Armitage 2010). Embedded dust grains are the seed for the enormous growth to bodies of planetary size. The first hint to the structure of protoplanetary discs has been provided by our own solar system. Through "smearing out" all planets and adding the missing fraction of volatile elements, one can estimate the structure and mass of the protoplanetary disc. Since the efficiency of planet formation is unknown, this yields only a lower limit – the minimum mass solar nebula (Hayashi 1981). The inferred surface density decreases with the distance from the sun as $\propto r^{-3/2}$.

Further insight has been obtained by the detection of an infrared excess of some stars, i. e. additional infrared radiation that does not originate from the star but an unresolved disc. Advancements in observation led in the mid-90s to the direct imaging of nearby star forming regions which opened a

⋆ E-mail: Glaschke@ari.uni-heidelberg.de (PG); Pau.Amaro-Seoane@aei.mpg.de (PAS, corresponding author); Spurzem@bao.ac.cn (RS)

new flourishing field in astronomy (see O'dell et al. 1993,for an example with Hubble Space Telescope images)). Since then a big amount of observations of protoplanetary discs across the electromagnetic range has been obtained and interpreted, both space-based (ISO, Spitzer and now Herschel) and in the ground (VLT, VLT-I, Subaru, Keck). We refer the reader to "The Star Formation Newsletter" URL[1] for an overview of the recent papers in star- and planet formation and the book review "Protostars and Planets V" (Reipurth et al. 2007).

Protoplanetary discs masses cover a range from 10^{-3} to $0.1 M_{\odot}$ with a peak around $0.01 M_{\odot}$ (data from Taurus/Ophiuchus Beckwith 1996), in accordance with mass estimates deduced from the minimum mass solar nebula. Since the disc lifetime can not be measured directly, it is derived indirectly from the age of young, naked (i. e. discless) stars which sets an upper limit. Pre-main sequence evolutionary tracks are used to gauge the stellar ages, providing lifetimes of a few 10^6 years. The subsequent evolution of the disc proceeds in several stages.

Two different scenarios have been proposed to explain the formation of kilometre-sized planetesimals.

(i) One process is the gravitational instability of the dust component in a protoplanetary disc that leads to the direct formation of larger bodies. Goldreich & Ward (1973) propose that an initial growth phase of dust grains leads to a thin dust disc that undergoes a gravitational collapse. As the dense dust layer decouples from the gas, it rotates with the local Keplerian velocity, whereas the gas component rotates slower as it is partially pressure supported. This gives rise to a velocity shear at the boundary, which may excite turbulence through the Kelvin-Helmholtz-instability. Since the motion of small dust grains in the boundary layer is coupled to the gas, the turbulent velocity field could suppress the formation of a stratified dust layer, which is a necessary prerequisite for the gravitational instability (Weidenschilling 1977).

(ii) The collisional agglomeration of dust particles is an opposed formation mechanism. Relative velocities are dominated by the Brownian motion in the early phases of the growth process. This mechanism becomes increasingly inefficient with growing mass, but successively the particles decouple from the gas and settle to the midplane – a process that yields even larger velocities with increasing mass. The sedimentation initiates a growth mode that is similar to the processes in rain clouds: Larger grains drop faster, thus accreting smaller grains on their way to the midplane. Turbulence may modify this basic growth scenario by forcing the dust grains in a convection-like motion. Dust grains still grow during the settling process, but the turbulent velocity field could mix up dust from the midplane, and a new cycle begins. Each cycle adds a new layer to the dust grains – a mechanism that also operates in hail clouds – until the grains are large enough to decouple from the turbulent motion. Again, turbulence plays an important role in determining the growth mode and the relative velocities. While the relative velocities are high enough to allow for a fast growth, it is not clear *a priory* that collisions are sticky enough to allow for a net growth. High speed encounters lead to fragmentation, which counteracts agglomeration (e.g. Blum & Wurm 2000,and references therein). An important bottleneck in the agglomeration process is the fast orbital decay of 1 m sized boulders. Their orbital lifetime is as short as 100 years, and a quick increase in size – at least over one order of magnitude – is needed to reduce the radial drift significantly.

To overcome the difficulties associated with each of these scenarios, modifications have been proposed. Magnetohydrodynamic simulations include electro-magnetic interactions in hydrodynamical calculations. See the reviews of Balbus & Hawley 1998 and Balbus. 2003. The MHD simulations by Johansen et al. (2006) show that trapping of larger particles in turbulent vortices helps in increasing the orbital lifetime, but could also trigger local instabilities that may lead to the direct formation of planetesimals (Inaba et al. 2005). Johansen et al. (2007) describe a gravoturbulence mechanism as a feasible pathway to planetesimal formation in accreting circumstellar discs.

The details of agglomeration have drawn a lot of attention and are still under question (see Kempf et al. 1999; Paszun & Dominik 2009; Wada et al. 2009), but the successive agglomeration of planetesimals is commonly accepted.

1.1 Formation of protoplanets

The further growth of planetesimals proceeds through mutual collisions, where the initial phase involves a large number of particles and is well described by a coagulation equation (Safronov 1969). While earlier works (e.g. Nakagawa et al. 1983) focussed mainly on the evolution of the size distribution, subsequent refinements of the evolution of the random velocities showed that it is important to evolve the size distribution and the velocity dispersion in a consistent way. A fixed velocity dispersion is an oversimplification, which changes the growth mode and increases the growth timescale as well (Wetherill 1989).

The initial growth is quite democratic. All planetesimals grow roughly at the same rate and the maximum of the size distribution is shifted gently towards larger sizes. As soon as gravitational focusing and dynamical friction become important, the growth mode changes to a qualitatively different mechanism. Efficient gravitational focusing leads to a growth timescale (which we denote as $M/\dot{M}$) that decreases with mass. Hence larger bodies grow faster than the smaller planetesimals, a trend that is further supported by energy equipartition due to planetesimal–planetesimal encounters. This dynamical friction keeps the largest bodies on nearly circular orbit, thus the relative velocities are small and gravitational focusing remains efficient. Smaller planetesimals are stirred up into eccentricity orbits, which slows down their growth rate compared to the largest bodies. This accelerated growth, denoted as runaway growth of planetesimals (see e.g Greenberg et al. 1978; Wetherill 1989, 1993), shortens the growth timescale to a few 10^5 years. The term runaway growth stresses that the growth timescale of a particle decreases with mass, hence the largest body "runs away" to the high mass end of the distribution (see Kokubo 1995).

While energy equipartition increases the velocity dispersion with decreasing mass of the planetesimals, addi-

[1] `http://www.ifa.hawaii.edu/users/reipurth/newsletter.htm`

tional damping due to the gaseous disc leads to a turn-over at smaller sizes. However, higher relative velocities may lead nevertheless to destructive encounters, but these fragmentation events could even speed up the growth (Chambers 2006; Bromley & Kenyon 2006, 2010). Since smaller bodies are more subjected to gas drag, their velocity dispersion is smaller which allows an efficient accretion by the runaway bodies. Moreover, smaller particles damp the velocity dispersion of the largest bodies more efficiently.

As runaway proceeds, the system becomes more and more dominated by few big bodies – the protoplanets . Due to the dominance of few, very large bodies one can not use statistical methods anymore to simulate the problem.

Kobayashi et al. (2010) derived analytical expressions for the final masses of these planetary "embryos", as they call them, including the role of planetesimal depletion due to collisional disruption. They conclude that the final mass in the minimum-mass solar nebula at several AU can achieve ~ 0.1 Earth mass within 10^7 years.

1.2 Oligarchic growth

The runaway growth of large bodies (i. e. protoplanets) ceases to be efficient as soon as the protoplanets start to control the velocity dispersion of the remaining planetesimals in their vicinity. Gravitational focusing becomes less effective, therefore the growth timescale increases with size and the growth mode changes to oligarchic growth. The protoplanets still grow faster than the field planetesimals[2], but the masses of the protoplanets remain comparable.

A combination of dynamical friction due to the field planetesimals and perturbations from the neighbouring protoplanets conserves a separation of five to ten Hill radii between neighbouring bodies. Therefore only planetesimals from a limited area, the feeding zone, are accreted by a given protoplanet. If this zone is emptied, they have reached their final isolation mass (Kokubo et al. 1998; Kokubo & Ida 2000, 2002).

As the damping of the remaining field planetesimals is weak enough, further growth is dominated by mutual perturbations among the protoplanets, which leads to giant impacts. Protoplanets beyond the "snow line" (or ice condensation point) can grow larger than 5–15 earth masses and initiate the formation of giant planets. If the protoplanetary disc is very massive in the inner planetary system, this may lead to an in-situ formation of hot Jupiters (see Bodenheimer et al. 2000).

Kokubo et al. (2006) used the oligarchic growth model of protoplanets to statistically quantify the giant impact stage (i.e. collisions of protoplanets to form planets) with $N-$body simulations. They found that for steeper surface density profiles, large planets usually form closer to the star.

1.3 Migration

The proposed three-stage scenario of planet formation covers the dominant growth processes, but a major mechanism is still missing – the migration of bodies in the system. Migration is a generic term that summarises a set of different mechanisms that lead to secular radial drift of bodies (see e.g. the review of Papaloizou 2006; Armitage 2010):

(i) The dissipation due to the remaining gas disc leads to an orbital decay of the planetesimals. While this poses a severe problem for 1 m–sized objects, larger bodies drift very slowly inward. One denotes this process as type 0 migration.

(ii) Planets which are embedded in a gaseous disc launch spiral density waves at the inner and outer Lindblad resonances, which leads to an exchange of angular momentum with the resonant excited waves. This type I migration leads to a robust inward migration independent of the density profile. The perturbation from the planet is small, hence linear perturbation theory is in principle applicable (Ward 1986), but recent work proves that non-linear and radiative effects can be quite important (Paardekooper et al. 2011).

(iii) If the protoplanet is massive enough, it opens a gap in the gaseous disc and excites waves through tidal interaction with the gaseous disc. An imbalance of the exchange of angular momentum with the inner and outer part of the disc leads to type II migration. The strong interaction between planet and disc leaves the linear regime and requires a numerical solution of the hydrodynamic equations (Lin & Papaloizou 1979).

(iv) Even with an opened gap, the planet still channels gas between the inner and outer part of the disc. While this corotational flow is to some extent already present during type II migration, it dominates the angular momentum balance in the case of type III migration due to an asymmetry in the leading and trailing part of the flow. An imbalance allows for an efficient exchange of angular momentum with this corotational gas stream, which gives rise to a remarkably fast migration (Masset & Papaloizou 2003). Type III migration is subjected to a positive feedback: A faster migration increases the asymmetry in the corotational flow, which speeds up the migration. Both inward and outward migration are possible.

The four migration types modify the three-stage scenario in different ways.
Type 0 migration is most efficient for small planetesimals (i. e. 1 km or smaller). It poses a severe problem during the early stages of planet formation, as it may induce a significant loss of solid material, accompanied by a global change in the initial surface density (Kornet et al. 2001). The importance of this process diminishes as planetesimal growth proceeds, but it still leads to the loss of collisional fragments during the final disc clearing.
Type I is most important for protoplanets (i. e. 0.1 earth masses or larger). It leads to an orbital decay of protoplanets, but this does not only imply a loss of protoplanets, but also breaks the conditions for isolation. Migrating bodies can accrete along their way through the disc and are thus not constrained to a fixed feeding zone, which may increase the isolation mass.
Type II and *type III* mainly influences giant planets (larger than 10 earth masses) causing an inward or outward migration, depending on the angular momentum exchange balance. It can explain the large number of giant planets close to their host star (hot Jupiters) found in extrasolar systems. An important issue is the timescale of the migration process.

[2] The term "field planetesimals" denotes in the following the smooth component of smaller planetesimals.

If the migration is too fast, virtually all planets spiral inward and leave an empty system behind.

Migration is a powerful process with the capability to reshape an entire planetary system, at least for gas giants, since terrestrial planets are likely not affected by migration in gaseous discs (Bromley & Kenyon 2011). However, it also requires some "parking mechanism" which terminates migration before all planets (or protoplanets) are lost. Inhomogeneities in the gaseous disc may change the crucial momentum balance of the inner and outer part of the disc, thus stalling or even reversing the drift of a planet.

The migration processes end after the dissipation of the gaseous disc due to photoevaporation or star-star encounters (i. e. after a few 10^6 to 10^7 years).

The formation of a planetary system is a vital process that is driven by the interplay between the different growth phases and the migration of planets and protoplanetary cores (i. e. the precursors of giant planets). While the preceding sections only summarised the main evolutionary processes, even more processes could influence the formation of planetary systems. A fast accretion of giant planets in the outer parts of a planetary system could introduce further perturbations on the inner part and may even trigger the formation of terrestrial planets. Moreover, the stellar environment in dense star clusters and multiple stellar systems also perturbs planet formation, which would require an even broader view on the problem. Last but not least, the excitation due to gravitational coupling to gas surface density fluctuations plays an important role (see e.g. Ida et al. 2008).

Any approach to planet formation can hardly include this wealth of different phenomena, thus it is important to focus on a well-defined subproblem. In this work we focus on the formation of protoplanets for the following reasons:

(i) The size, growth timescale and spacing of the protoplanets is a key element in the planet formation process.

(ii) The protoplanet growth is well-defined by different growth modes. It starts with the already formed planetesimals ($\approx$ 1–5 km) and ends, when orbital crossing of the protoplanets initiates the final growth phase.

(iii) The planetesimals are large enough to treat the remaining gaseous disc as a small perturbation.

(iv) The protoplanets are small enough to neglect tidal interaction with the disc in the inner planetary system. Collective planetesimal–protoplanet interaction are also negligible.

Though the protoplanet formation is a well-posed subproblem, our approach has to incorporate various mechanisms and techniques to cover the full size range of the problem (see the review of Armitage 2010,for a extensive review of the problem and references therein). However, it is still accessible to theoretical calculations to some extent which provide a safe ground for the analysis of the results.

In next section 2 we briefly summarise the analytical model that we use for our initial models. Then we introduce the problem that we will address later, as well as the timescales and quantities of interest for the general problem of planetesimal formation: In sections 3, 4, 5 we give a description for a test particle moving around a body with a central mass M, the Hill's problem and the protoplanet growth from a theoretical point of view, respectively.

Subsequently, in section 6 and its corresponding subsections, we ascertain the direct-summation part of the general algorithm, as well as the required modifications and additions to tailor it to the specific planetesimal problem. In section 9 we introduce our collisional treatment, including fragmentation, collisional cascades, migration and coagulation. The statistical part of the algorithm is described in detail in section 16. Finally, in section 17 we explain how to bring together both schemes into a single numerical tool and in section 18 we give our conclusions and compare our scheme to other numerical approaches.

2 INITIAL MODELS

The basis for all planet formation models is the structure of the protoplanetary disc. We summarise the pioneering work of Hayashi (1981) to have a robust initial model at hand. Subsequent evolution of the disc may change this simple approach, but it is still a valuable guideline.

A basic estimate of the minimum surface density of solid material in the disc can be deduced from the mass and location of the present planets in the solar system:

$$\Sigma_{\mathrm{solid}}(r) = \begin{cases} 7.1\,(r/1\mathrm{AU})^{-3/2}\,\mathrm{g/cm^2} & 0.35 \leqslant r \leqslant 2.7 \\ 30.0\,(r/1\mathrm{AU})^{-3/2}\,\mathrm{g/cm^2} & 2.70 \leqslant r \leqslant 36.0 \end{cases}$$

The discontinuity at 2.7 AU stems from the location of the ice condensation point (or snow line) that allows the formation of icy grains in the outer solar system. Furthermore, the total surface density is estimated through the chemical composition of the disc, which gives the ratio of gas to solids. A fiducial value is 1:0.017 (see Cameron 1973). The surface density of the gas component is therefore:

$$\Sigma_{\mathrm{gas}}(r) = 1700 \left(\frac{r}{1\mathrm{AU}}\right)^{-3/2} \frac{\mathrm{g}}{\mathrm{cm^2}} \tag{1}$$

Since the dust content is rather low, the gaseous component is transparent to the visible solar radiation. Thus the gas temperature follows from the radiation balance:

$$T = T_0 \left(\frac{r}{1\mathrm{AU}}\right)^{-1/2} \left(\frac{L}{L_\odot}\right)^{1/4} \qquad T_0 = 280\ \mathrm{K} \tag{2}$$

L is the solar luminosity during the early stages, normalised by the present value $L_\odot$. The three-dimensional density structure is given by an isothermal profile

$$\rho_{\mathrm{gas}}(r,z) = \frac{\Sigma_{\mathrm{gas}}}{\sqrt{2\pi}h} \exp\left(-\frac{z^2}{2h^2}\right) \tag{3}$$

with a radially changing scale height h:

$$\begin{aligned} h &= \frac{c_s}{\Omega} \qquad c_s = \left(\frac{k_B T}{\mu m_{\mathrm{H}}}\right)^{1/2} \qquad \mu = 2.34 \\ &= \frac{c_s^{(0)}}{\Omega_0} \left(\frac{r}{1\mathrm{AU}}\right)^{5/4} \left(\frac{L}{L_\odot}\right)^{1/8} \qquad c_s^{(0)} = 993.56 \frac{\mathrm{m}}{\mathrm{s}} \end{aligned} \tag{4}$$

c_s is the sound velocity of an ideal gas with a mean molecular weight μ in units of the hydrogen mass m_{H}. Since the density profile is related to a radially varying pressure, the gas velocity deviates from the local Keplerian velocity. The balance of forces relates the angular velocity Ω_g to the pressure gradient:

$$r\Omega_g^2 = r\Omega^2 + \frac{1}{\rho}\frac{dP}{dr} \tag{5}$$

Thus the angular velocity Ω_g of the gas is (see e.g. Adachi et al. 1976):

$$\begin{aligned}\Omega_g &= \Omega\sqrt{1-2\eta_g(r)}\\ \eta_g &= -\frac{1}{2}\frac{d\ln(\rho_{\rm gas}c_s^2)}{d\ln(r)}\left(\frac{c_s}{v_K}\right)^2\end{aligned} \tag{6}$$

It is more appropriate to formulate the rotation of the gaseous disc in terms of a velocity lag Δv_g normalised to the local Keplerian velocity v_K:

$$\begin{aligned}\Delta v_g &= r(\Omega_g-\Omega)\\ &\approx -\eta_g v_K\end{aligned} \tag{7}$$

A typical value of Δv_g for the minimum mass solar nebula at 1 AU is $\Delta v_g = -60$ m/s.

This simple model provides a brief description of the initial disc. However, the subsequent evolution further modifies the structure of the protoplanetary disc. Since embedded dust grains are coupled to the gas, it is likely that a global migration of solid material changes the surface density. Moreover, the dust grains are chemically processed, depending on the local temperature and composition which introduces additional spatial inhomogeneities. When the growing particles pass the critical size of ~ 1 metre, the strong onset of radial migration may lead to a final reshaping of the distribution of solid material. While these restrictions weaken the validity of this approach as the "true" initial model, it is still a robust guideline to choose reasonable surface densities for the solid and the gaseous component after the formation of planetesimals.

3 KEPLER ORBITS

Planetesimals in a protoplanetary disc are subjected to various perturbations: Close encounters change their orbits, a small but steady gas drag gives rise to a radial drift and accretion changes the mass of the planetesimals. While all theses processes drive the disc evolution on a timescale of at least a few thousand years, each planetesimal moves most of the time on an orbit close to an unperturbed Kepler ellipse. Though the protoplanetary disc introduces additional perturbations, the central potential dominates for typical disc masses around $0.01\,M_\odot$. Therefore the classical orbital elements still provide a proper framework to study planetesimal dynamics.

The orbital elements of a test particle moving around a mass M are:

$$\begin{aligned}E &= \frac{v^2}{2}-\frac{GM}{r}\\ a &= -\frac{GM}{2E}\\ e &= \sqrt{1-\frac{L^2}{GMa}}\\ \cos(i) &= \frac{L_z}{L}\end{aligned} \tag{8}$$

$$\begin{aligned}\mathbf{L} &= \mathbf{r}\times\mathbf{v}\\ e\cos(\phi_E) &= \frac{rv^2}{GM}-1\\ e\sin(\phi_E) &= \frac{\mathbf{r}\cdot\mathbf{v}}{\sqrt{GMa}}\end{aligned} \tag{9}$$

a is the semimajor axis, e is the eccentricity and i is the inclination of the orbit. As long as no dominant body is structuring the protoplanetary disc, it is justified to assume axisymmetry. Hence the argument of the perihelion ω, the longitude of the ascending node Ω and the eccentric anomaly ϕ_E are omitted in the statistical description.

The deviation of planetesimal orbits from a circle is quite small. Thus it is appropriate to expand the above set of equations. A planetesimal at a distance r_0, in a distance z above the midplane and with a velocity (v_r, v_ϕ, v_z) with respect to the local circular velocity v_K has orbital elements (leading order only):

$$a \approx r_0 + 2\frac{r_0 v_\phi}{v_K} \tag{10}$$

$$e^2 \approx \frac{v_r^2+4v_\phi^2}{v_K^2} \tag{11}$$

$$i^2 \approx \frac{z^2}{r_0^2}+\frac{v_z^2}{v_K^2} \tag{12}$$

These expressions allow us later a convenient transformation between the statistical representation through orbital elements and the utilisation of a velocity distribution function.

4 HILL'S PROBLEM

When two planetesimals pass close by each other, they exchange energy and angular momentum and separate with modified orbital elements. Successive encounters transfer energy between planetesimals with different masses, driving an evolution of the overall velocity distribution.

It seems that an encounter is a two-body problem, as there are only two planetesimals involved, but the central mass has also a major influence turning the problem into a three-body encounter [3] The complexity of the problem is considerably simplified by reducing it to Hill's problem (Hill 1878).

Consider two masses m_1 and m_2 that orbit a much larger mass M_c, where both masses are small compared to the central mass M_c. The mass ratio $m_1 : m_2$ could be arbitrary. This special type of a three body problem is denoted as Hill's problem, originally devised to calculate the orbit of the moon. It provides a convenient framework to examine planetesimal encounters in the potential of a star. The equations of motion[4] of the two planetesimals including the

[3] The term three-body encounter does not imply a close passage of all involved bodies, but emphasises the strong influence of a third one.

[4] The following derivation is quite common to the literature (see e.g. Ida 1990; Hénon 1986). The later work uses a slightly different scaling.

central potential and their mutual interaction are:

$$\ddot{\mathbf{r}}_1 = -\mathbf{r}_1 \frac{GM_c}{r_1^3} - (\mathbf{r}_1 - \mathbf{r}_2)\frac{Gm_2}{r_{12}^3}$$
$$\ddot{\mathbf{r}}_2 = -\mathbf{r}_2 \frac{GM_c}{r_2^3} - (\mathbf{r}_2 - \mathbf{r}_1)\frac{Gm_1}{r_{12}^3} \qquad (13)$$

We now introduce the relative vector $\mathbf{r}$ and the centre-of-mass $\mathbf{R}$:

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 \qquad \mathbf{R} = \frac{m_2\mathbf{r}_2 + m_1\mathbf{r}_1}{m_1 + m_2} \qquad (14)$$

Furthermore, the equations of motion are transformed to a corotating set of coordinates which are scaled by the mutual Hill radius $r_{\rm Hill}$ and the local Kepler frequency Ω

$$x = \frac{r' - a_0}{r_{\rm Hill}}$$
$$y = \frac{a_0(\phi - \Omega t)}{r_{\rm Hill}}$$
$$z = \frac{z'}{r_{\rm Hill}}$$
$$r_{\rm Hill} = a_0 \sqrt[3]{\frac{m_1 + m_2}{3M_c}} \qquad (15)$$

where (r', ϕ, z') are heliocentric cylindrical coordinates. a_0 is the radius of a properly chosen reference orbit, while the mutual Hill radius $r_{\rm Hill}$ is an intrinsic length scale of the problem : Since the two orbiting planetesimals are small compared to the central star, the Hill radius is much smaller than the size a_0 of the reference orbit. Hence it is possible to expand the central potential about the reference orbit. This yields approximate equations for the relative motion and the centre-of-mass:

$$\ddot{x} = 2\dot{y} + 3x - 3x/r^3$$
$$\ddot{y} = -2\dot{x} - 3y/r^3$$
$$\ddot{z} = -z - 3z/r^3 \qquad (16)$$

$$\ddot{X} = 2\dot{Y} + 3X$$
$$\ddot{Y} = -2\dot{X}$$
$$\ddot{Z} = -Z \qquad (17)$$

The equations 16 and 17 have some interesting properties: Firstly, the centre-of-mass motion separates from the interaction of the two bodies. Secondly, the scaled relative motion is independent of the masses m_1 and m_2, implying a fundamental similarity of planetesimal encounters[5]. As Eq. 17 is a simple linear differential equation, one readily obtains the solution of the centre-of-mass motion

$$X = b - e\cos(t - \tau)$$
$$Y = -\frac{3}{2}bt + \psi + 2e\sin(t - \tau)$$
$$Z = i\sin(t - \omega) \qquad (18)$$

which is equivalent to a first-order expansion of a Kepler ellipse. ω and τ are the longitudes of the ascending node and the pericentre, while e and i are the eccentricity and inclination scaled by the reduced (i. e. dimensionless) *Hill radius* $r_{\rm Hill}/a_0$. The value of b depends on the choice of the reference orbit, but it is natural to set $b = 0$ which implies that the centre-of-mass defines the reference orbit.

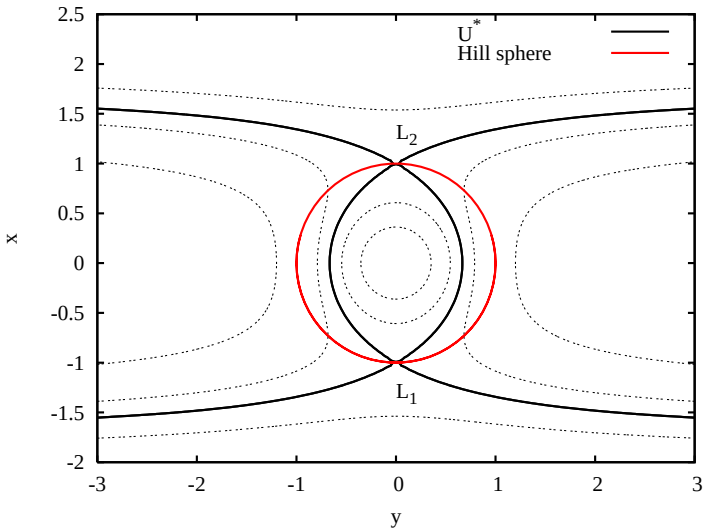

Figure 1. Equipotential lines for the effective potential U at $z = 0$ (see Eq. 20). $U = U^*$ refers to the largest allowed volume, which is enclosed by the Hill sphere and the two Lagrange points L_1 and L_2.

While the nonlinear nature of the relative motion (see Eq. 16) prevents any general analytical solution, Eq. 18 provides at least an asymptotic solution for a large separation of the planetesimals, where b can be interpreted as an impact parameter. Nevertheless, small b do not necessarily imply close encounters, as opposed to the standard definition of the impact parameter. However, the expression $b = a_2 - a_1$ provides a measure of the distance of the two colliding bodies without invoking the complicated encounter geometry. A special solution to Eq. 16 are the unstable equilibrium points L_1 and L_2 at $(x, y, z) = (\pm 1, 0, 0)$, denoted as Lagrange points[6]. In addition, an inspection of Eq. 16 reveals that the *Jacobi energy* E_J is conserved:

$$E_J = \frac{1}{2}\left(\dot{x}^2 + \dot{y}^2 + \dot{z}^2 + z^2 - 3x^2\right) - \frac{3}{r} \qquad (19)$$

Since the kinetic energy is always a positive quantity, the following inequality holds:

$$E_J \geqslant U = \frac{1}{2}\left(z^2 - 3x^2\right) - \frac{3}{r} \qquad (20)$$

Thus the allowed domain of the particle motion is enclosed by the equipotential surfaces of the effective potential U. A subset of these equipotential surfaces restricts the allowed domain to the vicinity of the origin (see Fig. 1). The largest of these surfaces passes through the Lagrange points L_1 and L_2. Hence we identify the Hill radius (or Hill sphere) as the maximum separation which allows the bound motion of two planetesimals[7].

[5] section 16 makes extensive use of this property.

[6] The additional Lagrange points L_3–L_5 are missing due to the Hill approximation.

[7] The same argument applies to the tidal boundary in cluster

Beside the numerical solution of the equations of motion, it is useful to define osculating (or instantaneous) orbital elements

$$\begin{aligned} b &= 4x + 2\dot{y} \\ i^2 &= z^2 + \dot{z}^2 \\ e^2 &= \dot{x}^2 + (3x + 2\dot{y})^2 \end{aligned} \tag{21}$$

$$\begin{aligned} \psi &= -2\dot{x} + y + \frac{3}{2}bt \\ \omega &= t - \arctan(z, \dot{z}) \\ \tau &= t - \arctan(\dot{x}, 3x + 2\dot{y}) \end{aligned} \tag{22}$$

$$E_J = \frac{1}{2}(e^2 + i^2) - \frac{3}{8}b^2 - \frac{3}{r} \tag{23}$$

which provide a convenient description of the initial relative orbit and the modified orbit after the encounter.

5 PROTOPLANET GROWTH

Our work follows the evolution of a planetesimal disc into few protoplanets, including the full set of interaction processes. Hence we summarise the main aspects of protoplanet growth first to provide a robust framework.

Although the sizes of the planetesimal cover a wide range, they virtually form two main groups: The smaller field planetesimals and the embedded protoplanets (or their precursors). This two-group approximation (e.g. Wetherill 1989; Ida 1993) allows one to have a clearer insight into the growth process.

During the initial phase all planetesimals share the same velocity dispersion independently of their mass. The initial random velocities are low enough for an efficient gravitational focusing. Hence, the growth rate of a protoplanet with mass M radius R can be estimated as (e.g. Ida et al. 1993):

$$\dot{M} \approx v_{\rm rel} \frac{\Sigma}{H} \pi R^2 \left(1 + \frac{2GM}{R v_{\rm rel}^2}\right) \tag{24}$$

Eq. 24 is the two-body accretion rate, which should be modified due to the gravity of the central star. Nevertheless this approximation gives an appropriate description to discuss the basic properties of the growth mode. The scale height H (Eq. 215) and the relative velocity $v_{\rm rel}$ are related to the mean eccentricity $e_m = \sqrt{\langle e^2 \rangle}$ of the field planetesimals:

$$H \approx v_{\rm rel}/\Omega \qquad v_{\rm rel} \approx e_m a \Omega \tag{25}$$

Thus the accretion rate in the limit of strong gravitational focusing ($2GM/R \gg v_{\rm rel}^2$) is:

$$\dot{M} \approx 2\pi R \frac{\Sigma G M}{a^2 \Omega e_m^2} \tag{26}$$

$$\propto M^{4/3} \tag{27}$$

If the protoplanets are massive enough, they start to control the velocity dispersion of the planetesimals in their vicinity.

dynamics or the Roche lobe in stellar dynamics, which are equivalent to the Hill sphere.

The width Δa of this sphere of influence, the heating zone, is related to the Hill radius $R_{\rm Hill}$ of the protoplanet (Ida 1993):

$$\begin{aligned} \Delta a = \Delta\tilde{a} R_{\rm Hill} &= 4 R_{\rm Hill} \sqrt{\frac{4}{3}(\tilde{e}_m^2 + \tilde{i}_m^2) + 12} \\ h &= \sqrt[3]{\frac{M}{3M_c}} \end{aligned} \tag{28}$$

$\tilde{e}_m$ and $\tilde{i}$ are eccentricity and inclination of the field planetesimals, scaled by the reduced Hill radius h of the protoplanet. M_c is the mass of the central star. The condition that the protoplanet controls the velocity dispersion of the field planetesimals reads (Ida et al. 1993):

$$\frac{2M^2}{2\pi a \Delta a} > \Sigma m \tag{29}$$

This condition is equivalent to a lower limit of the protoplanetary mass:

$$\frac{M}{m} > \left(\frac{\pi \Delta\tilde{a}}{\sqrt[3]{3}}\right)^{3/5} \left(\frac{\Sigma a^2}{M_c}\right)^{3/5} \left(\frac{m}{M_c}\right)^{3/5} \tag{30}$$

M/m depends on several parameters, but reasonable values yield $M/m \approx 50$–100. The velocity dispersion in the heated region is roughly

$$v \approx R_{\rm Hill} \Omega \tag{31}$$

which gives an interesting relation to the condition that leads to gap formation. A protoplanet can open a gap in the planetesimal component if it is larger than a critical mass $M_{\rm gap}$ (Rafikov 2001)

$$\frac{M_{\rm gap}}{M_c} \approx \begin{cases} \frac{\Sigma a^2}{M_c} \left(\frac{m}{M_c}\right)^{1/3} & \text{if} \quad v \lesssim \Omega r_{\rm Hill} \\ \frac{\Sigma a^2}{M_c} \left(\frac{m}{M_c}\right)^{1/3} \left(\frac{\Omega r_{\rm Hill}}{v}\right)^2 & \text{if} \quad v \gg \Omega r_{\rm Hill} \end{cases} \tag{32}$$

where $r_{\rm Hill}$ is the Hill radius of the field planetesimals. If the velocity dispersion v is controlled by the protoplanet, Eq. 32 together with Eq. 31 demonstrate that the condition for gap formation is equivalent to Eq. 30, i. e. the efficient heating of the field planetesimals implies gap formation and vice versa. The higher velocity dispersion of the field planetesimals (see Eq. 31) reduces the growth rate given by equation 26 to

$$\dot{M} \approx 6\pi \Sigma \Omega \frac{R R_{\rm Hill}}{\tilde{e}_m^2} \tag{33}$$

$$\propto M^{2/3} \tag{34}$$

Different mass accretion rates imply different growth mode. If two protoplanets have different masses M_1 and M_2, their mass ratio evolves as:

$$\frac{d}{dt}\frac{M_2}{M_1} = \frac{M_2}{M_1}\left(\frac{\dot{M}_2}{M_2} - \frac{\dot{M}_1}{M_1}\right) \tag{35}$$

When the growth timescale $M/\dot{M}$ decreases with mass, a small mass difference increases with time. This is the case for Eq. 26, which gives rise to runaway accretion. As soon as the protoplanets control the velocity dispersion of the field planetesimals, the growth timescale increases with mass and therefore the protoplanet masses become more similar as they grow oligarchically.

The field planetesimals also damp the excitation due

Σ[g/cm^2]	$M_{\rm iso}/M_\odot$	$M_{\rm iso}/M_\oplus$
2	3.91×10^{-8}	0.013
10	4.33×10^{-7}	0.144
100	1.37×10^{-5}	4.548

Table 1. Isolation mass for different surface densities at $r = 1$ AU and $M_c = 1\,M_\odot$.

to protoplanet–protoplanet interactions and keep them on nearly circular orbits. The balance between these scatterings and the dynamical friction due to smaller bodies establishes a roughly constant orbital separation b (Kokubo 1997):

$$b = R_{\rm Hill} \sqrt[5]{\frac{7\tilde{e}_m^2 M}{2\pi\Sigma a R}} \tag{36}$$

$$\tilde{b} = b/R_{\rm Hill} \tag{37}$$

R is the radius of the protoplanet, M is its mass and $\tilde{e}$ is the reduced eccentricity of the field planetesimals. The stabilised spacing prevents collisions between protoplanets, but it also restricts the feeding zone – the area from which a protoplanet accretes. If all matter in the feeding zone is accreted by the protoplanet, it reaches its final isolation mass (Kokubo & Ida 2000):

$$M_{\rm iso} = 2\pi b a \Sigma \tag{38}$$

Inserting Eq. 36 yields the isolation mass in units of the mass of the host star M_c:

$$\begin{aligned} M_{\rm iso}/M_c = (112\pi^4)^{3/8} \left(\frac{1}{3}\right)^{5/8} \left(\frac{4\pi}{3}\right)^{1/8} \left(\tilde{e}_m^2\right)^{3/8} \\ \left(\frac{a^2\Sigma}{M_c}\right)^{3/2} \left(\frac{a^3\rho}{M_c}\right)^{1/8} \approx 19.67 \times \left(\tilde{e}_m^2\right)^{3/8} \\ \left(\frac{a^2\Sigma}{M_c}\right)^{3/2} \left(\frac{a^3\rho}{M_c}\right)^{1/8} \end{aligned} \tag{39}$$

There is a weak dependence on the density ρ of the protoplanet, but the most important parameter is the surface density Σ. If we take the minimum mass solar nebula as an example, the radial dependence of the surface density implies an isolation mass that grows with increasing distance to the host start. Hence protoplanets beyond some critical radial distance are massive enough (larger than $\approx 15\,M_\oplus$ Bodenheimer 1986) to initiate gas accretion from the protoplanetary disc.

As the protoplanets approach the isolation mass, interactions with the gaseous disc and neighbouring protoplanets become increasingly important. We estimate the onset of orbit crossing by a comparison of the perturbation timescale $\tau_{\rm pert}$ of protoplanet–protoplanet interactions with the damping timescale $\tau_{\rm damp}$ due to planetesimal–protoplanet scatterings. Since the protoplanets are well separated ($\tilde{b} \approx 5 \ldots 10$), it is possible to apply perturbation theory (Petit 1986,see e.g.):

$$\tau_{\rm pert} \approx \frac{\tilde{b}^5}{7h\Omega} \tag{40}$$

We anticipate section 16 (see Eq. 221) to derive the damping timescale

$$\tau_{\rm damp} \approx \frac{1}{2} \frac{T_r^{3/2}}{\sqrt{2\pi} G^2 \ln(\Lambda)(M+m) n_0 m} \tag{41}$$

where T_r and T_z are the radial and vertical velocity dispersion of the field planetesimals. Hence the criterion for the onset of orbital crossing is:

$$\tau_{\rm pert} < \tau_{\rm damp} \tag{42}$$

As the protoplanets control the velocity dispersion of the field planetesimals (see Eq. 31), this condition reduces to:

$$\begin{aligned} \Sigma_M &> \Sigma_m \ln(\Lambda) \frac{72}{7\pi} \left(\frac{\tilde{b}}{\tilde{e}}\right)^4 \\ &> \Sigma_m \times f \end{aligned} \tag{43}$$

Thus orbital crossing sets in when the mean surface density Σ_M of the protoplanets exceeds some fraction f of the field planetesimal density Σ_m. While the factor f depends strongly on the separation $\tilde{b}$ of the protoplanets, a fiducial value is $f \approx 1$, in agreement with the estimates of Goldreich et al. (2004).

The onset of migration and the resonant interaction of protoplanets with the disc and other protoplanets terminates the local nature of the protoplanet accretion process and requires a global evolution of the planetary system. While the final stage deserves a careful analysis, further research is beyond the scope of this work.

6 DIRECT-SUMMATION TECHNIQUES FROM THE STANDPOINT OF PLANETARY DYNAMICS: FIRST STEPS

The protoplanet formation is essentially an N–body problem. Although we seek for a more elaborated solution to this problem which benefits from statistical methods, the pure N–body approach is a logical starting point. Direct calculations with a few thousand bodies have provided us with a valuable insight into the growth mode (Ida 1992; Kokubo 1996,see e.g.), but they are also powerful guidelines that help developing other techniques. Statistical calculations rely on a number of approximations and "exact" N–body calculations provide the necessary, unbiased validation of the derived formula.

The choice of the integrator is a key element in the numerical solution of the equations of motion. Our requirements are the stable long-term integration of a few ten thousand planetesimals with the capability of treating close encounters, collisions and the perspective to evolve it into an improved hybrid code. Approximative methods like the Fast Multipole Method or Tree codes have a scaling of the computational time close to N, but the accuracy in this regime is too poor to guarantee the stable integration of Keplerian orbits (compare the discussion in Hernquist et al. 1993; Spurzem 1999).

The class of exact methods (all scale asymptotically with N^2) roughly divides in two parts:

(i) Symplectic methods (see e.g. Wisdom 1991 or the SYMBA code, Duncan et al. 1998) rely on a careful expansion of the Hamiltonian which guarantees that the numerical integration follows a perturbed Hamiltonian. While there is

still an integration error, all properties of a Hamiltonian system like conservation of phase-space volume are conserved by the numerical integration. The drawback of these very elegant methods is that the symplecticity is immediately broken by adaptive time steps, collisions or complicated external forces if no special precaution is taken. Even the numerical truncation error breaks the symplecticity to some extend (Skeel 1999). Moore & Quillen (2010) recently developed a parallel integrator for graphics processing unit (GPU) that uses symplectic and Hermite algorithms according to the resolution needed. The algorithm is similar to SYMBA, but less accurate.

(ii) The second group represents the "classical" methods that are based on Taylor expansions of the solution. They come in different flavours like implicit methods, predictor–corrector integrators or iterated schemes. Time symmetric methods stand out among these different approaches, as they show no secular drift in the energy error. These integrators are so well-behaved that one may even call them "nearly symplectic".

Taking all the requirements into account we have chosen NBODY6++[8], an integrator which is the most recent descendant from the famous N–body code family from S. Aarseths' "factory". This version was parallelised by Spurzem (1999), which opened the use of current supercomputers.

This parallel version, named NBODY6++, offers many versatile features that were included over the past years and more and more refined as time passed by. While all these elaborated tools deserve attention, we restrict ourselves for brevity to the components which contribute to the planetesimal problem. The main components of the code are:

(i) Individual time steps and a block time step scheme.
(ii) Ahmad–Cohen neighbour scheme.
(iii) Hermite scheme.
(iv) KS–Regularisation for close two-body encounters.

We will explain each of these features and highlight their advantages for the main goal of this work. This is important to understand the immediate first level of modifications required to adapt the numerical scheme to planetary dynamics. We present later, in section 8, the next level of complexity in the modifications to be carried out to mold the numerical tool to our purposes.

6.1 Individual Time Steps

The choice of the time step controls the accuracy as well as the efficiency of any given integrator. Too small time steps slow down the integration without necessity, whereas too large values increase the error. An efficient solution to this dilemma is an adaptive time step that is adjusted after each integration step according to a specified accuracy limit. While the idea is quite clear, there is no unique receipt how to choose the proper time step. A common approach for N–body systems is to use the parameters at hand (like particle velocity, force, etc.) to derive a timescale of the particle motion. This procedure leaves enough space for a wealth of different time step criteria. NBODY6++ uses the standard Aarseth expression (Aarseth 1985)

$$\Delta t = \sqrt{\eta \frac{F F^{(2)} + (F^{(1)})^2}{F^{(1)} F^{(3)} + (F^{(2)})^2}} \tag{44}$$

which makes use of the force and the time derivatives up to third order.

The time step choice is not unique in multi-particle systems. One solution is to take the minimum of all these values as a shared (or global) time step, but this is not recommended unless all individual steps are of the same size.

The second option is to evolve each particle track with its own, individual time step. This method abandons the convenient synchronisation provided by a global step, therefore each integration of a particle demands a synchronisation of all particles through predictions. Since the prediction of all particles is $\mathcal{O}(N)$, it is counter-balanced by the saving of force computations. Nevertheless the overhead is still significant, so a further optimisation might be desired. The basic idea of the block time step method is to force particles in groups that are integrated together, which reduces the number of necessary predictions by a factor comparable to the mean group size. These groups are enforced through two constraints. The first condition is a discretisation of the steps in powers of two:

$$\Delta t = 2^{-k} \qquad k \geqslant 0 \tag{45}$$

This condition increases the chances that two different particles share the same timelevel, but it also reduces roundoff errors since the time steps are now exactly representable numbers. The second condition locks the "phases" of particles with the same time step

$$T_i \equiv 0 \bmod \Delta t_i \tag{46}$$

i. e. the particle time T_i is an integer multiple of the actual time step Δt_i so that all particles with the same step share the same block. Note that a step can not be increased at any time T_i, but only when the second condition (Eq. 46) for the larger step is met.

6.2 Ahmad–Cohen Neighbour Scheme

The first N–body codes calculated always the full force (i. e. summation over all particles) to integrate a particle. But not all particles contribute with the same weight to the total force. Distant particles provide a smooth, slowly varying regular force, whereas the neighbouring particles form a rapidly changing environment which gives rise to an irregular force. The Ahmad-Cohen neighbour scheme (Ahmad & Cohen 1973) takes advantage of this spatial hierarchy by dividing the surrounding particles in the already mentioned two groups according to the neighbour sphere radius R_s. Both partial forces fluctuate on different timescales, which are calculated according to Eq. 44. The key to the efficiency of the method is the inequality

$$\Delta t_{\rm irr} \ll \Delta t_{\rm reg} \tag{47}$$

Regular forces are extrapolated between two full force calculations

$$\mathbf{F}_{\rm reg} = \mathbf{F}^{(0)}_{\rm reg} + \Delta t\, \mathbf{F}^{(1)}_{\rm reg} + \frac{1}{2}(\Delta t)^2 \mathbf{F}^{(2)}_{\rm reg} + \frac{1}{6}(\Delta t)^3 \mathbf{F}^{(3)}_{\rm reg} \tag{48}$$

[8] Aarseth (1999) gives a nice review on the remarkable history of the NBODY-codes, more details are given in Aarseth (2003).

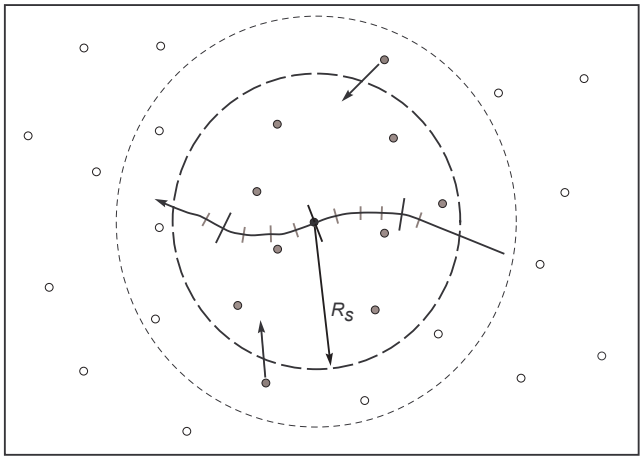

Figure 2. All particles inside the neighbour sphere R_s are selected as neighbours (filled circles). As the neighbour list is fixed during regular steps (marked by black lines), a second shell includes possible intruders.

with a third order accurate expression, whereas the plain NBODY6++ uses only linear extrapolation (see the next sections for a detailed discussion).

6.3 Hermite Scheme

The Hermite scheme is a fourth-order accurate integrator that was applied first by Makino (1992) to the integration of a planetesimal system. This scheme is used to integrate single particles and CM-bodies[9] in the main part of NBODY6++. It is accomplished by a predictor and a corrector step. The prediction is second order accurate:

$$\begin{aligned}\mathbf{x}_p &= \mathbf{x}_0 + \mathbf{v}_0\Delta t + \frac{1}{2}\mathbf{a}_0(\Delta t)^2 + \frac{1}{6}\dot{\mathbf{a}}_0(\Delta t)^3 \\ \mathbf{v}_p &= \mathbf{v}_0 + \mathbf{a}_0\Delta t + \frac{1}{2}\dot{\mathbf{a}}_0(\Delta t)^2 \end{aligned} \tag{49}$$

Now the acceleration is evaluated at the predicted position to derive the second and third order time derivatives of the force:

$$\mathbf{a}_0^{(2)} = \frac{-4\dot{\mathbf{a}}_0 - 2\dot{\mathbf{a}}_p}{\Delta t} + \frac{-6\mathbf{a}_0 + 6\mathbf{a}_p}{\Delta t^2} \tag{50}$$

$$\mathbf{a}_0^{(3)} = \frac{6\dot{\mathbf{a}}_0 + 6\dot{\mathbf{a}}_p}{\Delta t^2} + \frac{12\mathbf{a}_0 - 12\mathbf{a}_p}{\Delta t^3} \tag{51}$$

Using the additional derivatives one can improve the prediction:

$$\begin{aligned}\mathbf{x}_c &= \mathbf{x}_p + \frac{1}{24}\mathbf{a}_0^{(2)}(\Delta t)^4 + \frac{1}{120}\mathbf{a}_0^{(3)}(\Delta t)^5 + \mathcal{O}(\Delta t^6) \\ &= \mathbf{x}_p + \left(-\frac{3}{20}\mathbf{a}_0 + \frac{3}{20}\mathbf{a}_p\right)(\Delta t)^2 - \left(\frac{7}{60}\dot{\mathbf{a}}_0 + \frac{1}{30}\dot{\mathbf{a}}_p\right) \\ &(\Delta t)^3 + \mathcal{O}(\Delta t^6) \end{aligned} \tag{52}$$

$$\begin{aligned}\mathbf{v}_c &= \mathbf{v}_p + \frac{1}{6}\mathbf{a}_0^{(2)}(\Delta t)^3 + \frac{1}{24}\mathbf{a}_0^{(3)}(\Delta t)^4 + \mathcal{O}(\Delta t^5) \\ &= \mathbf{v}_p + \frac{1}{2}(-\mathbf{a_0} + \mathbf{a_p})\Delta t + \left(-\frac{5}{12}\dot{\mathbf{a}}_0 - \frac{1}{12}\dot{\mathbf{a}}_p\right)(\Delta t)^2 + \\ &\mathcal{O}(\Delta t^5) \end{aligned} \tag{53}$$

[9] CM denotes centre–of–mass, see the section on KS-Regularisation for more details.

The corrected positions are fourth order accurate. While the Hermite scheme is robust and stable, even in combination with the neighbour scheme, it is not accurate enough to integrate planetesimal orbits efficiently.

6.4 Hermite Iteration

The Hermite scheme bears the potential for a powerful extension, since it is a predictor–corrector scheme. An essential part of this scheme is the calculation of the new forces with the predicted positions, but it should improve the accuracy if they are recalculated using the corrected positions. The new forces are readily used to improve the corrected values, which closes the scheme to an iteration loop – the Hermite iteration (Kokubo et al. 1998).

There are only few modifications necessary to obtain the iterated scheme. The particle prediction remains second-order accurate:

$$\begin{aligned}\mathbf{x}_p &= \mathbf{x}_0 + \mathbf{v}_0\Delta t + \frac{1}{2}\mathbf{a}_0(\Delta t)^2 + \frac{1}{6}\dot{\mathbf{a}}_0(\Delta t)^3 \\ \mathbf{v}_p &= \mathbf{v}_0 + \mathbf{a}_0\Delta t + \frac{1}{2}\dot{\mathbf{a}}_0(\Delta t)^2 \end{aligned} \tag{54}$$

The force and its first time derivative are calculated to derive higher derivatives according to Eq. 50 and 51:

$$\mathbf{a}_p = \mathbf{f}(\mathbf{x}_p, \mathbf{v}_p) \tag{55}$$

$$\dot{\mathbf{a}}_p = \mathbf{f}(\mathbf{x}_p, \mathbf{v}_p) \tag{56}$$

The new corrector omits the highest order term in the position, making the velocity and the position to the same order accurate:

$$\begin{aligned}\mathbf{x}_c &= \mathbf{x}_p + \frac{1}{24}\mathbf{a}_0^{(2)}(\Delta t)^4 + \mathcal{O}(\Delta t^5) \\ &= \mathbf{x}_p + \frac{1}{4}(\mathbf{a}_p - \mathbf{a}_0)(\Delta t)^2 + \left(-\frac{1}{6}\dot{\mathbf{a}}_0 - \frac{1}{12}\dot{\mathbf{a}}_p\right) \\ &(\Delta t)^3 + \mathcal{O}(\Delta t^5) \end{aligned} \tag{57}$$

$$\begin{aligned}\mathbf{v}_c &= \mathbf{v}_p + \frac{1}{2}(-\mathbf{a}_0 + \mathbf{a}_p)\Delta t + \left(-\frac{5}{12}\dot{\mathbf{a}}_0 - \frac{1}{12}\dot{\mathbf{a}}_p\right) \\ &(\Delta t)^2 + \mathcal{O}(\Delta t^5) \end{aligned} \tag{58}$$

It seems unreasonable to drop one order in the position, but a reformulation of the predictor–corrector step reveals that this slight change yields a time symmetric scheme:

$$\begin{aligned}\mathbf{v}_c &= \mathbf{v}_0 + \frac{1}{2}(\mathbf{a}_p + \mathbf{a}_0)\Delta t - \frac{1}{12}(\dot{\mathbf{a}}_p - \dot{\mathbf{a}}_0)(\Delta t)^2 \\ \mathbf{x}_c &= \mathbf{x}_0 + \frac{1}{2}(\mathbf{v}_c + \mathbf{v}_0)\Delta t - \frac{1}{12}(\mathbf{a}_p - \mathbf{a}_0)(\Delta t)^2 \end{aligned} \tag{59}$$

The iteration is achieved by returning to Eq. 55–56 with the predicted positions replaced by the more accurate values $\mathbf{x}_c, \mathbf{v}_c$. Although the integration does not need the second and third time derivative of the forces explicitly, it is useful to provide them at the end of the iteration for the calculation of the new time step:

$$\mathbf{a}^{(2)}(t + \Delta t) = \frac{2\dot{\mathbf{a}}_0 + 4\dot{\mathbf{a}}_p}{\Delta t} + \frac{6\mathbf{a}_0 - 6\mathbf{a}_p}{\Delta t^2} \tag{60}$$

$$\mathbf{a}^{(3)}(t + \Delta t) = \frac{6\dot{\mathbf{a}}_0 + 6\dot{\mathbf{a}}_p}{\Delta t^2} + \frac{12\mathbf{a}_0 - 12\mathbf{a}_p}{\Delta t^3} \tag{61}$$

Numerical tests show that convergence is reached after one or two iterations, making this approach very efficient. It needs less force evaluations than the Hermite scheme for

the same accuracy. Moreover, its time symmetry suppresses a secular drift of the energy error.

6.5 Extended Hermite Scheme

The Hermite scheme is an integral part of NBODY6++ and proved its value in many applications. It would have been natural to improve the performance with an additional iteration, but our first tentative implementations showed rather negative results: The iterated scheme was more unstable, slower and even less accurate than the plain Hermite scheme. An inspection of the code structure revealed that the Ahmad–Cohen neighbour scheme is the cause of this.

Each particle integration is composed of two parts – frequent neighbour force evaluations and less frequent total force evaluations including derivative corrections. Every regular correction leads to an additional change in the position of a particle, which introduces a spurious discontinuity in the neighbour force and its derivatives. The Hermite iteration reacts to this artificial jump in two ways: It increases the regular correction, and – what is more important – it amplifies any spurious error during the iteration which leads to an extreme unstable behaviour.

Since the Hermite iteration is a key element in the efficient integration of planetesimal orbits, we sought for a modification of the Hermite scheme that circumvents the depicted instability. The problem gives already an indication of a possible solution. A scheme with much smaller corrections would not suffer from the feedback of spurious errors.

NBODY6++ stores already the second and third time derivative of the forces for the time step calculation. A manifest application of these derivatives at hand is the improvement of the predictions to fourth order:

$$\mathbf{x}_p = \mathbf{x}_0 + \mathbf{v}_0\Delta t + \frac{1}{2}\mathbf{a}_0(\Delta t)^2 + \frac{1}{6}\dot{\mathbf{a}}_0(\Delta t)^3 + \frac{1}{24}\mathbf{a}_0^{(2)}(\Delta t)^4 + \frac{1}{120}\mathbf{a}_0^{(3)}(\Delta t)^5 \quad (62)$$

$$\mathbf{v}_p = \mathbf{v}_0 + \mathbf{a}_0\Delta t + \frac{1}{2}\dot{\mathbf{a}}_0(\Delta t)^2 + \frac{1}{6}\mathbf{a}_0^{(2)}(\Delta t)^3 + \frac{1}{24}\mathbf{a}_0^{(3)}(\Delta t)^4 \quad (63)$$

The prediction to fourth order was used in the iterative schemes of Kokubo et al. (1998); Mikkola & Aarseth (1998). Again, the new forces $\mathbf{a}_p$ and $\dot{\mathbf{a}}_p$ are calculated to improve $\mathbf{x}_p$ and $\mathbf{v}_p$ – but with a modified corrector:

$$\mathbf{a}_p = \mathbf{f}(\mathbf{x}_p, \mathbf{v}_p) \quad (64)$$

$$\dot{\mathbf{a}}_p = \mathbf{f}(\mathbf{x}_p, \mathbf{v}_p) \quad (65)$$

$$\mathbf{a}_n^{(2)}(t) = \frac{-4\dot{\mathbf{a}}_0 - 2\dot{\mathbf{a}}_p}{\Delta t} + \frac{-6\mathbf{a}_0 + 6\mathbf{a}_p}{\Delta t^2} \quad (66)$$

$$\mathbf{a}_n^{(3)}(t) = \frac{6\dot{\mathbf{a}}_0 + 6\dot{\mathbf{a}}_p}{\Delta t^2} + \frac{12\mathbf{a}_0 - 12\mathbf{a}_p}{\Delta t^3} \quad (67)$$

$$\mathbf{x}_c = \mathbf{x}_p + \frac{1}{24}(\mathbf{a}_n^{(2)} - \mathbf{a}_0{}^{(2)})(\Delta t)^4 + \frac{1}{120}(\mathbf{a}_n^{(3)} - \mathbf{a}_0{}^{(3)})(\Delta t)^5 \quad (68)$$

$$\mathbf{v}_c = \mathbf{v}_p + \frac{1}{6}(\mathbf{a}_n^{(2)} - \mathbf{a_0}^{(2)})(\Delta t)^3 + \frac{1}{24}(\mathbf{a}_n^{(3)} - \mathbf{a_0}^{(3)})(\Delta t)^4 \quad (69)$$

Finally, the derivatives are updated:

$$\mathbf{a}_0(t + \Delta t) = \mathbf{a}_p \quad (70)$$

$$\dot{\mathbf{a}}_0(t + \Delta t) = \dot{\mathbf{a}}_p \quad (71)$$

$$\mathbf{a_0}^{(2)}(t + \Delta t) = \mathbf{a}_n^{(2)} + \Delta t\, \mathbf{a}_n^{(3)} \quad (72)$$

$$\mathbf{a_0}^{(3)}(t + \Delta t) = \mathbf{a}_n^{(3)} \quad (73)$$

The new scheme has an appealing property, which is related to the usage of the higher force derivatives. As the predictor is fourth-order accurate, it is equivalent to one full Hermite step. Since the corrector uses new forces to improve the two highest orders, it is equivalent to a first iteration step. Thus we obtained a one-fold iterated Hermite scheme at no extra cost. This extended Hermite scheme reduces the energy error by three orders of magnitude, compared to the plain NBODY6++ with the same number of force evaluations.

6.6 KS–Regularisation

Two bodies undergoing a close encounter are integrated in a special set of regular coordinates that separates the relative motion from the motion of the centre-of-mass. A close encounter poses a numerical problem due to the singularity of the gravitational forces at zero separation. While the growing force amplifies roundoff errors as the two bodies approach each other closely, the collision is only an apparent singularity since the analytic solution stays well-defined. This opens the possibility of a proper coordinate transformation which removes the singularity from the equations of motion. The Kustaanheimo–Stiefel regularisation takes advantage of a four-dimensional set of variables to transform the Kepler problem into a harmonic oscillator (Kustaanheimo & Stiefel 1965). Perturbations are readily included in the new set of equations of motion.

The centre-of-mass is added as a pseudo-particle, the CM-body, which is integrated as a normal particle plus a perturbation force due to the deviation from a point mass. See Mikkola (1997) or Mikkola & Aarseth (1998) for more details.

7 ADDITIONAL FORCES FOR PLANETESIMAL DISC DYNAMICS

NBODY6++ only includes the gravitational interaction of all particles, therefore additional forces have to be added "per hand". A planetesimal disc requires two new forces: The presence of a central star introduces an additional central potential, while the gaseous component of the protoplanetary disc is the source of a friction force. It is important that the new forces are properly included in the neighbour scheme to assure that regular steps remain larger than irregular steps. Since a dissipative force breaks the energy conservation, one has to integrate the energy loss as well to maintain a valid energy error control. In the next subsections we describe how we have done this.

7.1 Central Potential

A star is much heavier than a planetesimal. Thus, the central star is introduced as a spatially fixed Keplerian potential:

$$\Phi_c = -\frac{GM_c}{r} \qquad \mathbf{F} = -\frac{GmM_c\,\mathbf{x}}{|\mathbf{x}|^3} \tag{74}$$

Since the orbital motion of the planetesimals sets the dominant (and largest) dynamical timescale in the system, we included the central force as a component of the regular force. Moreover, the central potential also introduces a strong synchronisation, since planetesimals in a narrow ring share virtually the same regular block time step.

7.2 Drag Force

As the whole planetesimal system is embedded in a dilute gaseous disc, each planetesimal is subjected to a small, but noticeable drag force. The drag regime[10] depends on the gas density and the size of the planetesimals. Kilometer-sized planetesimals are subjected to the deceleration

$$\frac{d\mathbf{v}}{dt} = -\frac{\pi C_D}{2m}\rho_{\rm gas}R^2|\mathbf{v}-\mathbf{v}_g|(\mathbf{v}-\mathbf{v}_g) \tag{75}$$

$$C_D = 0.5 \tag{76}$$

which is inversely proportional to the radius $R(m)$ of the planetesimal in this drag regime. v_g is the rotational velocity of the gaseous disc, which rotates slower than the planetesimal system as it is partially pressure supported. The drag force leads to an orbital decay $\dot{a}$ of the semimajor axis of a planetesimals:

$$\dot{a} = -\frac{3}{4}C_D\frac{\rho_{\rm gas}}{\rho_{\rm Body}}\frac{\langle(\Delta v)^2\rangle}{R(m)\Omega} \tag{77}$$

Thus smaller particles migrate faster, with a maximum at $R \approx 1$ m. Even smaller bodies couple to the gas, which reduces the effective drag force. The dissipation rate and its time derivative are:

$$\begin{aligned} W_{\rm drag} &= \mathbf{F}_{\rm drag}\cdot\mathbf{v} \\ \dot{W}_{\rm drag} &= \dot{\mathbf{F}}_{\rm drag}\cdot\mathbf{v} + \mathbf{F}_{\rm drag}\cdot\mathbf{F}_{\rm tot} \end{aligned} \tag{78}$$

We integrate the dissipation rate $W_{\rm drag}$ to maintain a valid energy error:

$$\Delta E = \int_{t_1}^{t_2} W_{\rm drag}\,dt \tag{79}$$

$$\Delta t = t_2 - t_1 \tag{80}$$

$$\begin{aligned} \Delta E = &\frac{1}{2}\left(W_{\rm drag,1}+W_{\rm drag,2}\right)\Delta t \\ &+\frac{1}{12}\left(\dot{W}_{\rm drag,1}-\dot{W}_{\rm drag,2}\right)\Delta t^2 + \mathcal{O}(\Delta t^5) \end{aligned} \tag{81}$$

The expression is fourth order accurate in accordance with the order of the extended Hermite scheme.

[10] The main drag regimes are Stokes (laminar flow), Epstein (mean free molecular path larger than object size) and Newton's drag law (turbulent flow). Weidenschilling (1977) provides a nice review on the different drag regimes.

7.3 Accurate integration of close encounters: Tidal perturbations of KS–Pairs and impact of the gaseous disc

Both new forces also demand a modification of the regularisation treatment. They perturb the relative motion of a KS–pair and modify the orbit of the centre-of-mass. While the modification of the equations of motion is rather clear, the neighbour scheme requires some additional work.
Let $\mathbf{r}_1, \mathbf{r}_2$ be the positions of the two regularised particles. The equations of motion read ($G = 1$)

$$\begin{aligned} \mathbf{r} &= \mathbf{r}_2 - \mathbf{r}_1 \\ \ddot{\mathbf{r}}_1 &= -M_c\frac{\mathbf{r}_1}{r_1^3} + m_2\frac{\mathbf{r}}{r^3} + \mathbf{F}_{\rm drag,1} \\ \ddot{\mathbf{r}}_2 &= -M_c\frac{\mathbf{r}_2}{r_2^3} - m_1\frac{\mathbf{r}}{r^3} + \mathbf{F}_{\rm drag,2} \end{aligned} \tag{82}$$

where the perturbations by other particles have been omitted for clarity. Centre-of-mass motion and the orbital motion are separated:

$$\ddot{\mathbf{r}} = -M\frac{\mathbf{r}}{r^3} - M_c\frac{\mathbf{r}_2}{r_2^3} + M_c\frac{\mathbf{r}_1}{r_1^3} + \mathbf{F}_{\rm drag,2} - \mathbf{F}_{\rm drag,1} \tag{83}$$

$$\ddot{\mathbf{R}} = -M_c\frac{m_1}{M}\frac{\mathbf{r}_1}{r_1^3} - M_c\frac{m_2}{M}\frac{\mathbf{r}_2}{r_2^3} + \frac{m_1}{M}\mathbf{F}_{\rm drag,1} + \frac{m_2}{M}\mathbf{F}_{\rm drag,2} \tag{84}$$

$$\mathbf{r} = \mathbf{r}_2 - \mathbf{r}_1 \qquad M = m_1 + m_2 \tag{85}$$

$$\mathbf{R} = \frac{1}{M}\left(m_1\mathbf{r}_1 + m_2\mathbf{r}_2\right) \tag{86}$$

Two new contributions show up due to the external forces: The KS–pair is tidally perturbed by the central star and influenced by the gaseous disc. While the aerodynamic properties of a single particle are well understood, two bodies revolving about each other may induce complex gas flows in their vicinity, which could invalidate the linear combination of the drag forces on each component. Therefore we drop the drag force term to avoid spurious dissipation. Since the dynamic environment allows virtually no stable binaries[11] in a planetesimal disc, the influence of the drag force on the encounter dynamics is negligible.

We further decompose the additional acceleration of the centre-of-mass motion, since the neighbour scheme benefits from a clear separation of the timescales. Therefore, the tidal perturbation is split into a smooth mean force and a perturbation force:

$$\ddot{\mathbf{R}} = \mathbf{F}_{\rm mean} + \mathbf{F}_{\rm pert}$$

$$\mathbf{F}_{\rm mean} = -M_c\frac{\mathbf{R}}{R^3}$$

$$\mathbf{F}_{\rm pert} = M_c\frac{\mathbf{R}}{R^3} - M_c\frac{m_1}{M}\frac{\mathbf{r}_1}{r_1^3} - M_c\frac{m_2}{M}\frac{\mathbf{r}_2}{r_2^3} \tag{87}$$

$$= 0 + \mathcal{O}(r^2) \tag{88}$$

The mean forces varies on the orbital timescale and is hence

[11] Tidal capturing of moons starts in the late stages of planet formation, but is limited to the planets or their precursors. However, the quiescent conditions in an early Kuiper belt lead to a more prominent role of binaries. See the summary of Astakhov et al. (2005).

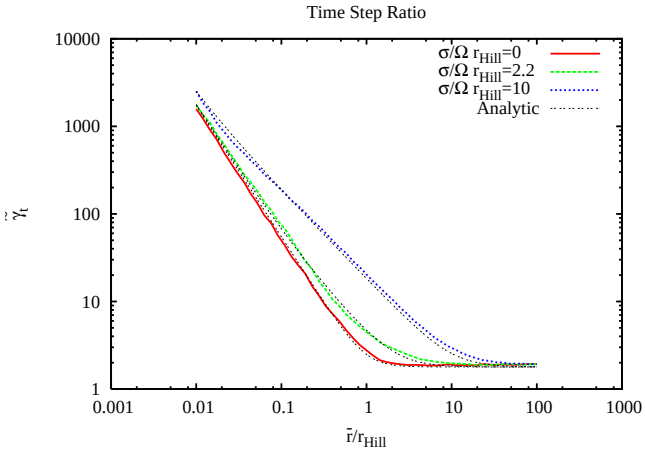

Figure 3. Time step ratio for $N_{\rm nb} = 100$. Curves are plotted for different values of $\sigma_v/(r_{\rm Hill}\Omega)$. The dotted line is approximation Eq. 95.

included as a regular force component, while the perturbation is treated as an irregular force as it changes with the internal orbital period of the pair.

8 FURTHER TUNING THE DIAL: OPTIMISING N–BODY TO THE DISC

An astrophysical simulation is a tool to analyse problems and predict dynamical systems which are not accessible to experiments. The design of a new simulation tool does not only require the careful implementation of the invoked physics, but also an analysis of the code performance to make best use of the available hardware.

We want to apply NBODY6++ for the first time to planet formation, a subject that is quite different to stellar clusters. The central star forces the planetesimals on regular orbits which need higher accuracy than the motion of stars in a cluster. In addition, the orbital motion also introduces a strong synchronisation among the planetesimals, thus allowing a more efficient integration.

We examine the differences due to the integration of a disc system in the following sections. In particular, we will address in detail the role of the geometry of the problem and the neighbour scheme, the prediction of the number of neighbouring particles, the communication, the block size distribution and the optimal neighbour particle number for the direct-summation of the massive particles in the protoplanetary system

8.1 Disc Geometry and Neighbour Scheme

The introduction of the neighbour scheme by Ahmad & Cohen (1973) has provided us with a technique to save a considerable amount of computational time in star cluster simulations. Since the average ratio of the regular to the irregular time step γ_t is of the order of 10, the integration is speeded up by the same factor. One may expect a similar speedup for planetesimal systems, but in this case the time step ratio is roughly three. The time step is calculated with the standard Aarseth time step criterion (it should be mentioned, however, that the relation of regular and irregular costs is more complicated with GPU technology)

$$\Delta t = \sqrt{\eta \frac{FF^{(2)} + (F^{(1)})^2}{F^{(1)}F^{(3)} + (F^{(2)})^2}} \tag{89}$$

where $F^{(i)}$ are the force and its time derivatives. It is applied to the calculation of the regular step using the regular force and accordingly to the irregular step based on the irregular force. The regular time step of a particle orbiting the central mass M_c at a distance r_0 is:

$$\Delta t_r = \sqrt{\eta_r}\frac{1}{\Omega} \qquad \Omega = \sqrt{\frac{GM_c}{r_0^3}} \tag{90}$$

For simplicity, we introduce the scaled timestep ratio $\tilde{\gamma}_t = \gamma_t\sqrt{\eta_{\rm irr}/\eta_{\rm reg}}$. The free parameters of the problem are the mean particle distance $\bar{r}$, the velocity dispersion σ_v (additional to the Keplerian shear), the particle mass m_i and the neighbour number $N_{\rm nb}$. We employ Hill's approximation for the central potential and obtain:

$$\Delta t_i \approx \frac{\sqrt{\eta_i}}{\Omega} f(\Omega, r_{\rm Hill}, \bar{r}, \sigma_v, N_{\rm nb}) \tag{91}$$

$$r_{\rm Hill} = r_0 \sqrt[3]{\frac{2m_i}{3M_c}} \tag{92}$$

f is a yet unknown function. Dimensional analysis leads to

$$\Delta t_i \approx \frac{\sqrt{\eta_i}}{\Omega} f\left(\frac{\sigma_v}{r_{\rm Hill}\Omega}, \frac{\bar{r}}{r_{\rm Hill}}, N_{\rm nb}\right) \tag{93}$$

$$\tilde{\gamma}_t \approx f\left(\frac{\sigma_v}{r_{\rm Hill}\Omega}, \frac{\bar{r}}{r_{\rm Hill}}, N_{\rm nb}\right) \tag{94}$$

which shows that the time step ratio is essentially controlled by the interparticle distance and the velocity dispersion. We generated different random realisations of planetesimals discs with different densities and velocity dispersions to cover the range of possible values. The neighbour number is fixed to $N_{\rm nb} = 100$ to reduce the noise due to small number statistics, but γ_t converges to a value independently of the neighbour number already for $N_{\rm nb} > 10$. Fig. 3 shows the numerical calculation of the time step ratio for various values of $\bar{r}$ and σ_v. A good approximation to the calculated values of $\tilde{\gamma}_t$ is:

$$\tilde{\gamma}_t \approx 1.79 \times \sqrt{1 + 1.03\frac{\sigma_v^2}{\bar{r}^2\Omega^2} + 0.94\frac{r_{\rm Hill}^3}{\bar{r}^3}} \tag{95}$$

Planetesimal discs have usually a small velocity dispersion (compared to the orbital velocity) and a low density in terms of the Hill radius, which leaves a major influence to the Keplerian shear. Since the shear motion is directly linked to the local Keplerian frequency, this synchronisation reduces γ_t to values smaller than ten. The numerical calculations show larger time step ratios with increasing velocity dispersion and for high densities[12], but planetesimal discs are far from these extreme parameter values.

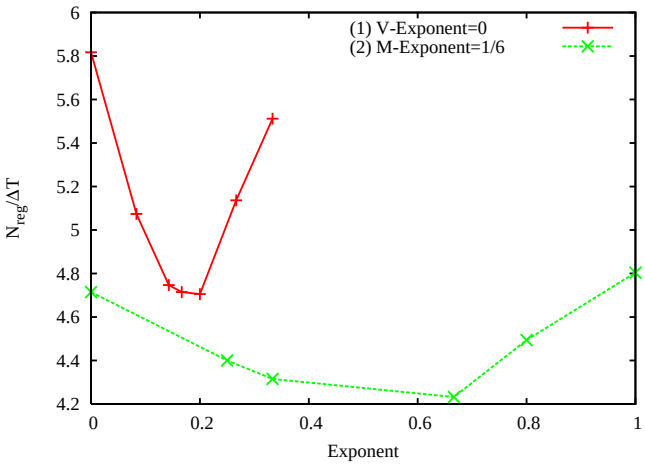

Figure 4. Regular steps per particle and per 1 N–body time in the inner core ($R_i < 0.5$) of a 5000 particle plummer model. Plotted are (1) different mass exponents with velocity exponent 0 and (2) different velocity exponents with mass exponent 1/6.

8.2 Optimal Neighbour Criterion

The standard neighbour criterion uses the geometrical distance: Particles are neighbours if their distance to the reference particle is smaller than a limit R_s. This criterion is simple and probably the best choice for an equal mass system. However, a multi-mass system may require a different criterion, since a massive particle outside the neighbour sphere could have a stronger influence than lighter particles inside the neighbour sphere. Also, relative effects are smaller at large distances. A more appropriate selection should rely on some "perturbation strength" of a particle.

It turned out that a better criterion is the magnitude of the fourth time derivative of the pairwise force $F_{ij}^{(4)}$, i. e. those particles are selected as neighbours which produce the largest integration error in accordance with the Hermite scheme. $F_{ij}^{(4)}$ is a complicated expression (compare Appendix A), but the leading term can be estimated via dimensional analysis:

$$F_{ij}^{(4)} \propto \frac{m_j v_{ij}^4}{r_{ij}^6} \tag{96}$$

We use this expression to define a new apparent distance between the integrated particle i and a neighbour j:

$$r_{\rm app} = r_{ij} \left(\frac{m_i}{m_j}\right)^{1/6} \left(\frac{v_s}{v_{ij}}\right)^{2/3} \tag{97}$$

v_s is an arbitrary scaling velocity to obtain a distance with dimension length. This new distance definition moves massive or fast neighbours to an apparently smaller distance, thus enforcing that these particles are preferentially included in the neighbour list. In addition, the modified distance is readily included in the conventional neighbour scheme. We tested different mass and velocity exponents to verify that Eq. 97 is the optimal choice. Figure 4 shows that these exponents are indeed the optimal choice for a Plummer model (Plummer 1911) with mass spectrum. The new scheme saves 25% of the force evaluations in the core, but the impact on a planetesimals system is smaller, as it is the case for the neighbour scheme. While a velocity dependent distance reduces the number of necessary full force evaluations, it introduces a distance changing with time which destabilises the integration. The result is a much larger energy error compared to the achieved speedup. Therefore we only recommend the mass modification of the apparent distance.

8.3 Neighbour Changes

The rate at which the neighbours of a given particle change has a noticeable influence on the accuracy of the code. During the course of an integration the second and third time derivative of the regular and irregular force are calculated from an interpolation formula (see Eq. 50 and Eq. 51). Whenever a particle leaves (or enters) the neighbour sphere, these derivatives are corrected by analytic expressions[13]. Hence many neighbour changes lead to a pronounced spurious difference.

We estimated the rate at which particles cross the neighbour sphere boundary to quantify this effect. Neighbour changes are due to the Keplerian shear and the superimposed random velocities of the particles. The two effects lead to:

$$\begin{aligned} N^{+/-} &= \Delta t_r \frac{N_{\rm nb}}{T_{\rm orb}} && \text{Shear} \\ N^{+/-} &= \frac{3}{2} \Delta t_r N_{\rm nb} \frac{\sigma_v}{R_{\rm nb}} \\ &= \Delta t_r \frac{N_{\rm nb}}{T_{\rm orb}} \frac{3\pi\sigma_v}{R_{\rm nb}\Omega} && \text{Dispersion} \end{aligned} \tag{98}$$

$R_{\rm nb}$ and $N_{\rm nb}$ are the neighbour sphere radius and the neighbour number, respectively. In practice, the neighbour changes due to the shear account for up to 80% of the total neighbour changes. The standard regular time step $\Delta t_r = 2^{-5}$ and 50 neighbours yield a change of one particle per regular step, which is fairly safe.

8.4 Neighbour Prediction

Each integration step is preceded by the prediction of all neighbours of the particles that are due. A regular step requires the full prediction of all particles, so there is no possibility to save computing time. In contrast, an irregular step calculates only neighbour forces, which requires the prediction of less particles. Thus the prediction of all particles to prepare an irregular step is a simple, but, depending on the block size, computational costly solution. It seems to be more efficient to predict only the required particles, but random access to the particle data and the complete check of all neighbour list entries introduces an additional overhead. Therefore large block sizes should favour the first approach, whereas the second approach is more suitable for small block sizes.

Both regimes are separated by a critical block size $N_{\rm irr}^*$. If $N_{\rm irr}$ particles with $\langle N_{\rm nb} \rangle$ neighbours are due, then only $N_{\rm merge}$ particles need to be predicted:

[12] $\bar{r}/r_{\rm Hill} < 1$ corresponds to an unstable self-gravitating disc.

[13] Appendix A gives a complete set of the force derivatives up to third order.

Process	0	1	2	3	4	5	6	7
Send to	1	2	3	4	5	6	7	0
Receive from	7	0	1	2	3	4	5	6

Table 2. Ring Communication. Communication partners are fixed, while the exchanged data varies. $n_p - 1$ cycles are needed.

Cycle	Process	0	1	2	3	4	5	6	7
1	Exchange with	1	0	3	2	5	4	7	6
2	Exchange with	2	3	0	1	6	7	4	5
3	Exchange with	4	5	6	7	0	1	2	3

Table 3. Hierarchical Communication. Communication partners change after every cycle. The exchanged data amount doubles with every new cycle, hence only $\ln_2(n_p)$ cycles are needed.

$$N_{\rm merge} \approx N_{\rm tot}\left(1 - \exp\left(-\frac{N_{\rm irr}\langle N_{\rm nb}\rangle}{N_{\rm tot}}\right)\right) \leqslant N_{\rm irr}\langle N_{\rm nb}\rangle \tag{99}$$

The size $N_{\rm merge}$ of the merged neighbour lists is smaller than the total number of neighbour list entries, since some particles are by chance members of more than one neighbour list. Performance measurements show that the prediction of the merged neighbour lists is 10 % more costly (per particle) than the full prediction, mainly due to additional sorting and a random memory access. Thus $N_{\rm irr}^*$ satisfies:

$$N_{\rm tot} = 1.1 \times N_{\rm merge} \tag{100}$$

Inserting Eq. 99 yields the critical block size:

$$N_{\rm irr}^* \approx 2.4\frac{N_{\rm tot}}{\langle N_{\rm nb}\rangle} \tag{101}$$

The prediction mode is chosen according to the actual block size.

8.5 Communication Scheme

NBODY6++ is parallelised using a copy algorithm. A complete copy of the particle data is located on each node, so the integration step of one particle does not need any communication. Therefore a block of $N_{\rm bl}$ particles is divided in n_p parts (n_p is the processor number), which are integrated by different processors in parallel. The integration step is completed by an all–to–all communication of the different subblocks to synchronise the particle data on all nodes. Hence the amount of communicated data is proportional to $N_{\rm bl} \times n_p$. A communication in a ring-like fashion (see table 2) needs $n_p - 1$ communication cycles, but a hierarchical scheme (see table 3) sends the same amount of data with only $\ln_2(n_p)$ communication cycles. The difference between the two approaches remains small, as long as the communication is bandwidth limited, i. e. the blocks are large. Small block sizes shift the bottleneck to the latency, which is significantly reduced by the second scheme – especially if the code runs on many processors.

Block	n_p	$\alpha[\mu s]$	$\tau_l[\mu s]$	A	B
Irregular	10	0.35	51	145	4.5
Regular	10	0.22	113	512	40
Irregular	20	0.35	308	877	8.8
Regular	20	0.22	368	1668	75

Table 4. Timings on a Beowulf cluster (Hydra, see table 5). See text for an explanation of the variables. Timings are obtained for a maximal neighbour number LMAX=64. In practice, B is twice as large due to storage rearrangements in NBODY6++. See Appendix 5 for details on the computers.

Block	n_p	$\alpha[\mu s]$	$\tau_l[\mu s]$	A	B
Irregular	8	0.29	255	837	1.7
Regular	8	0.60	981	1763	4.1
Irregular	16	0.28	188	700	1.7
Regular	16	0.60	306	561	6.7
Irregular	64	0.27	241	887	7.7
Regular	64	0.46	401	871	21.7

Table 6. Timings on the IBM. More than 32 processors require more than one node.

A hierarchical scheme reduces the latency, but nevertheless it is possible that the parallel integration is actually slower than a single CPU integration. We estimated both the runtime on one CPU and on a parallel machine to explore the transition between these two regimes. The latency time τ_l per communication is included in the wallclock time expressions for one regular/irregular step:

$$\tau_l = \alpha A$$
$$t_{\rm single} = \alpha N_{\rm bl} N_{\rm nb} \tag{102}$$
$$t_{\rm par} = \alpha\Big(\underbrace{\frac{N_{\rm bl}N_{\rm nb}}{n_p}}_{\rm Arithmetic} + \underbrace{A\ln_2(n_p)}_{\rm Latency} + \underbrace{BN_{\rm bl}}_{\rm Communication}\Big) \tag{103}$$

If $t_{\rm single}$ (runtime on a single CPU) is equal to $t_{\rm par}$ (parallel computation), one can deduce the critical block size $N_{\rm min}$ which gives the minimal block size for efficient parallelisation:

$$t_{\rm single} = t_{\rm par}$$
$$N_{\rm min} = \frac{A\ln_2(n_p)n_p}{N_{\rm nb}(n_p - 1) - Bn_p} \tag{104}$$

The hierarchical communication gives a minimal block size that increases logarithmically with the processor number. Eq. 103 gives immediately the speedup S and the optimal processor number for a certain block size $N_{\rm bl}$:

$$S = \frac{n_p}{1 + An_p\frac{\ln_2(n_p)}{N_{\rm bl}N_{\rm nb}} + B\frac{n_p}{N_{\rm nb}}}$$
$$n_{p,{\rm opt}}(N_{\rm bl}) = \frac{\ln(2)N_{\rm bl} \times N_{\rm nb}}{A} \quad - \text{Hierarchical} \tag{105}$$

A comparison to the optimal processor number for a ring

Name	Beowulf (Hydra)	Titan	JUMP
Institute	ARI/ZAH	ARI/ZAH	Forschungszentrum Jülich
Location	Heidelberg	Heidelberg	Jülich
Processors	20	64	1248
Speed	2.2GHz	3.2GHz	1.7GHz
Processors/Node	2	2	32
Network	Myrinet	Infiniband	Gigabit–Ethernet
Bandwidth	2Gbit/sec	20 Gbit/sec	10 Gbit/sec

Table 5. Specs of the different supercomputers used in running the algorithm

communication

$$n_{p,\mathrm{opt}}(N_{\mathrm{bl}}) = \sqrt{\frac{N_{\mathrm{bl}} \times N_{\mathrm{nb}}}{A}} \quad - \mathrm{Ring}$$
$$t_{\mathrm{par}} = \alpha\Big(\frac{N_{\mathrm{bl}} N_{\mathrm{nb}}}{n_p} + A n_p + B N_{\mathrm{bl}}\Big) \qquad (106)$$

stresses the efficiency of the hierarchical communication, since it allows a much larger processor number for a given problem size. Equation 102 and 103 are also useful to derive the total wallclock time, since the total runtime scales with the number of regular and irregular blocks:

$$N_{\mathrm{reg}} \approx T \frac{N^{1/3}}{N_{\mathrm{nb}}^{1/3}\sqrt{\eta_{\mathrm{reg}}}} \qquad (107)$$

$$N_{\mathrm{irr}} \approx T \frac{N^{1/3}}{\sqrt{\eta_{\mathrm{irr}}}} \qquad (108)$$

These equations are only approximate expressions, but they give the right order of magnitude without detailed calculations that need a precise knowledge of the N–body model. Table 4 and table 6 summarise the timing parameters drawn from our experience with the Hydra and JUMP parallel supercomputers.

8.6 Block Size Distribution

The preceding section showed that the block size is closely related to the efficiency of the parallelisation. Small blocks are dominated by the latency and the parallelisation could be even slower than a single CPU calculation. Therefore we derive the block size distribution for the block time step scheme to asses its influence on the efficiency.
Suppose that the time steps[14] h of all N particles in the model are distributed according to some known function f:

$$dp = f(N,h)dh \qquad (109)$$

f is in most cases a complicated function. It involves spatial averaging and integration over the velocity distribution, which could be quite complicated even for simple time step formulæ. Nevertheless there is a constraint on the time step distribution, simply because every particle has a neighbour within a finite distance: There is some upper limit $h_{\max}$ that

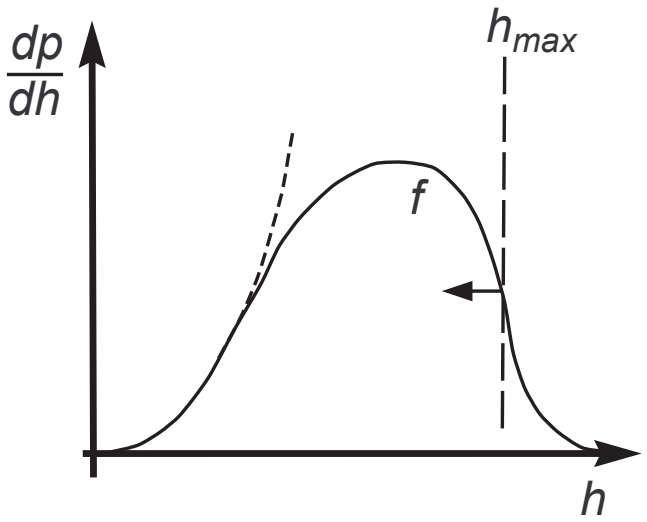

Figure 5. Timestep distribution $f = dp/dh$. The short-dashed line on the left indicates approximation Eq. 110, whereas the dashed line on the right defines a reasonable upper limit $h_{\max}$.

restricts the major fraction of the time steps to a finite interval. Thus it is possible to capture the main features of the time step distribution with an expansion around $h = 0$ (Fig. 5 sketches this approximation):

$$f \approx C(N)h^a \qquad h \leqslant h_{\max} \qquad (110)$$

a is the lowest non-vanishing order of the expansion. Now we consider a block level with the largest possible time step h_k. The number of particles N_{bl} in this block is:

$$N_{\mathrm{bl}} = N\int_0^{h_k} f dh \approx \frac{C(N)}{a+1}(h_k)^{a+1} \qquad (111)$$

According to the block time step scheme the number of blocks per time with the largest possible time step h_k is proportional to $(h_k)^{-1}$. Therefore the probability that a block size is in the range $[N_{\mathrm{bl}}, N_{\mathrm{bl}} + dN_{\mathrm{bl}}]$ is

$$dp \propto \sum_k \delta\left(N_{\mathrm{bl}} - \frac{C(N)}{a+1}h_k^{a+1}\right)\frac{1}{h_k}dN_{\mathrm{bl}} \qquad (112)$$

where δ is Dirac's delta function. The sum over the logarithmically equidistant time steps h_k is approximated by an

[14] We use h instead of Δt in this section to avoid unclear notation.

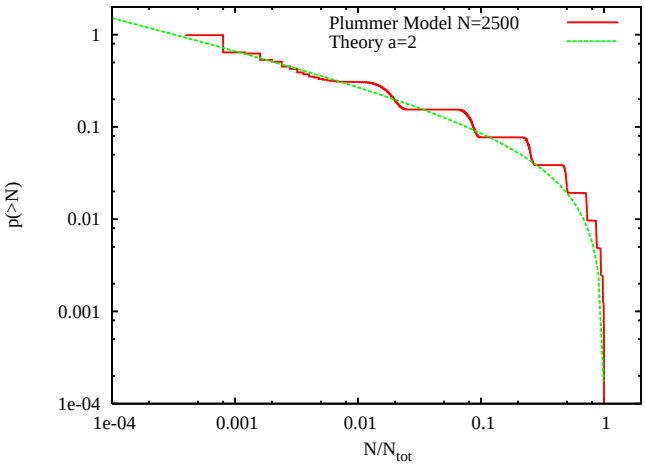

Figure 6. Cumulative irregular block size distribution for a $N = 2500$ particle Plummer model.

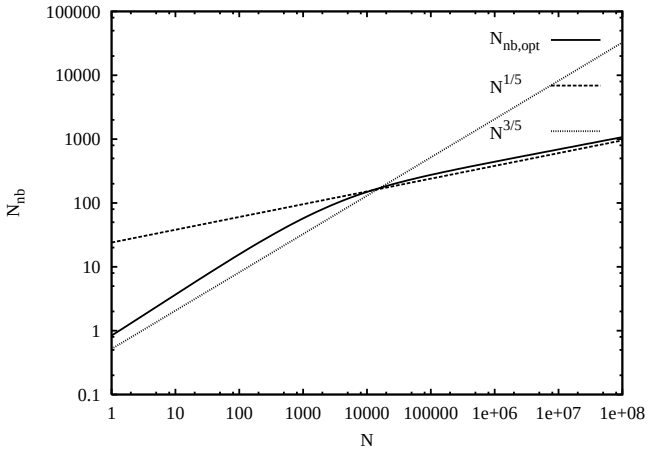

Figure 7. Optimal neighbour number as function of particle number N. The plot includes the numerical solution of Eq. 124 and the two asymptotic solutions. Timing constants are taken from a Beowulf cluster.

integral:

$$dp \propto \int_0^\infty \delta\left(N_{\rm bl} - \frac{C(N)}{a+1}h^{a+1}\right)\frac{d\ln(h)}{h}dN_{\rm bl}$$
$$\approx \frac{1}{a+1}N_{\rm bl}^{-(a+2)/(a+1)}dN_{\rm bl} \qquad (113)$$

Thus the average block size and the median of the block size distribution are:

$$\langle N_{\rm bl}\rangle \approx \frac{1}{a}N^{a/(a+1)}$$
$$\text{median}(N_{\rm bl}) \approx 2^{a+1} \qquad (114)$$

Special expressions for the average block size were already derived by Makino (1988), but the general relation of the time step distribution to the block size distribution is a new result. The median is surprisingly independent of the particle number, i. e. 50 % of all blocks are always smaller than a fixed value. It seems that this is a threat to the efficiency of the method, but the median of the wallclock time

$$\text{median}(N_{\rm bl}^2) \approx \frac{N}{2^{(a+1)/a}} \qquad (115)$$

demonstrates that these small blocks account only for a small fraction of the total CPU time. We confirmed the derived block size distribution (Eq. 113) by numerical calculations (see Fig. 6). The order parameter a is roughly two in (at least locally) homogenous systems, while an additional Keplerian potential reduces the order to $a = 1$. A planetesimal disc – or more precisely, a narrow ring of planetesimals – has a very narrow distribution of time steps since all particles share nearly the same orbital period. Thus the regular block size is always equal to the total particle number making the parallelisation very efficient.

8.7 Optimal Neighbour Number

We treated the mean neighbour number $N_{\rm nb}$ so far as some fixed value. But it is also a mean to optimise the speed of the integration. Large neighbour spheres reduce fluctuations in the regular forces allowing larger regular steps, which reduces the total number of force evaluations. But larger neighbour lists also imply a larger communication overhead, as all the neighbour lists have to be broadcast to synchronise the different nodes. The best choice balances these two extremes, thus maximising the speed.

Before we derive the optimal neighbour number on a parallel machine, we briefly summarise the known solution for a single CPU run (Makino 1988) for an extensive derivation. The computational effort of the irregular steps is proportional to the neighbour number, while the number of force evaluations for the regular steps is proportional to the total number of particles, reduced by the time step ratio γ_t:

$$\gamma_t := \frac{\Delta t_{\rm reg}}{\Delta t_{\rm irr}}$$
$$T_{\rm CPU} = f(N)\left(N_{\rm nb} + \frac{N}{\gamma_t(N_{\rm nb})}\right) \qquad (116)$$
$$\gamma_t(N_{\rm nb}) \approx N_{\rm nb}^{1/3} \qquad (117)$$

$f(N)$ collects all factors depending only on the total number of particles. Optimisation with respect to the neighbour number $N_{\rm nb}$ yields the well known result:

$$0 = \frac{d}{dN_{\rm nb}}T_{\rm CPU}$$
$$N_{\rm nb,opt} \propto N^{3/4} \qquad (118)$$

The calculation of the elapsed time for NBODY6++ on a PC cluster includes more terms. For clearness, we restrict ourselves to a rather simple model that involves only the dominant terms to show how parallelisation influences the optimal neighbour number. We make the following approximations:

(i) We only take the force calculation and communication into account.
(ii) We use the same time constants for regular and irregular expressions.
(iii) We neglect all numerical factors that are comparable to unity.

The total CPU time is an extension of Eq. 106, which is applied to the regular and the irregular step. A new constant B_n includes the neighbour list communication separately,

while all factors depending on N are represented by $f(N)$:

$$\begin{aligned} N_{\rm bl} &\approx N^{2/3} \\ \gamma_t &\approx N_{\rm nb}^{1/3} \\ T_{\rm irr} &= f(N)\left(\frac{N_{\rm bl}N_{\rm nb}}{p} + Ap + BN_{\rm bl}\right) \\ T_{\rm reg} &= \frac{1}{\gamma_t}f(N)\left(\frac{N_{\rm bl}N}{p} + Ap + (B + B_n N_{\rm nb})N_{\rm bl}\right) \\ T_{\rm tot} &= T_{\rm irr} + T_{\rm reg} \end{aligned} \tag{119}$$

Optimisation with respect to the processor number p leads to:

$$\begin{aligned} 0 &= \frac{\partial}{\partial p}T_{\rm tot} \\ p_{\rm opt} &= \sqrt{\frac{N^{2/3}N_{\rm nb}^{4/3} + N^{5/3}}{A(N_{\rm nb}^{1/3}+1)}} \\ &\approx \frac{N^{5/6}}{\sqrt{A}N_{\rm nb}^{1/6}} \end{aligned} \tag{120}$$

Further optimisation with respect to the neighbour number gives the expression:

$$0 = \frac{\partial}{\partial N_{\rm nb}}T_{\rm tot} \tag{121}$$

$$\begin{aligned} 0 = N_{\rm nb}^{4/3} - \frac{1}{3}N - \frac{1}{3}AN^{-2/3}p^2 - \frac{1}{3}Bp + \\ \frac{2}{3}N_{\rm nb}B_n p \end{aligned} \tag{122}$$

For a fixed p or $B_n = 0$ (very fast neighbour list communication), we recover for large N:

$$N_{\rm nb,opt} \propto N^{3/4} \tag{123}$$

In general, one can not neglect the neighbour list communication. Therefore we seek for the optimal choice of p and $N_{\rm nb}$, thus combining Eq. 120 and Eq. 122:

$$\begin{aligned} AN_{\rm nb}^{5/3} + \left(\frac{2}{3}N_{\rm nb} - \frac{B}{3B_n}\right)B_n\sqrt{A}N_{\rm nb}^{1/6}N^{5/6} = \\ \frac{1}{3}(N_{\rm nb}^{1/3}+1)NA \end{aligned} \tag{124}$$

Since this equation has no closed solution, we identify the dominant terms in Eq. 124 to calculate the asymptotic solution for large N:

$$N_{\rm nb,opt} \approx \left(\frac{A}{4B_n^2}\right)^{3/5}N^{1/5} \qquad N \gg \left(\frac{3A}{4B_n^2}\right)^{3/2} \tag{125}$$

For small N we get the approximated solution:

$$N_{\rm nb,opt} \approx \left(\frac{N}{3}\right)^{3/5} \qquad N < \left(\frac{3A}{4B_n^2}\right)^{3/2} \tag{126}$$

Fig. 7 compares the approximate expressions with the numerical solution of equation 124. In spite of the complicated structure of Eq. 124, both approximate expressions are reliable solutions. The example uses timing constants derived form our local Beowulf cluster:

$$A \approx 200 \quad B \approx 5 \quad B_n \approx 0.5 \tag{127}$$

If we compare the new optimal neighbour number to the single CPU expression (Eq. 118), we find that the influence of the neighbour list communication favours much smaller neighbour numbers. $N_{\rm nb}$ increases so slowly with the particle number that a neighbour number around 100 is a safe choice.

9 COLLISIONAL AND FRAGMENTATION MODEL

The growth of planetesimals proceeds through collisions among planetesimals which form (at least in a sufficient fraction of incidents) larger bodies with a net gain of accreted matter. But some collisions are mere destructive events that shatter and disperse the colliding planetesimals. Small bodies are more susceptible to destruction, but they are also driven to high relative velocities due to the global energy equipartition making them even more vulnerable. A model that attempts to cover the full size range from objects ranging between a kilometer and the size of Mars needs a realistic collision algorithm that covers both fragmentation and accretion. Some examples in the literature are Cazenave et al. (1982); Beauge & Aarseth (1990). In our algorithm we use the approach of Glaschke (2003), which was applied to asteroid families. The following section describes the approach to fragmentation that is implemented in the code.

9.1 Handy quantities for quantifying the models

Two colliding bodies are equal in the sense that their intrinsic properties are not different. Only the comparison of two bodies defines the larger body – usually denoted as target – and the smaller one denoted as *projectile*. The two terms stem from laboratory experiments where they indicate much more than different sizes. A small projectile is shot on a target at rest to study the various parameters related to fragmentation. In the following, projectile and target only indicate the relative size of the two bodies.

The collision of two bodies initiates a sequence of complex phenomena. Shock waves run through the material, flaws start to grow rapidly breaking the bodies apart in many pieces. Some kinetic energy is transferred to the fragments, which leads to the ejection of fragments at different velocities in various directions. If the fragment cloud is massive enough, some of the larger fragments may capture debris. This post-collisional accretion is denoted as reaccumulation.

Although the depicted scenario is quite complex, there are a few measures that describe the most important aspects:

(i) Mass of the largest fragment M_L, or dimensionless $f_l = M_L/M$ where M is the combined mass of the two colliding bodies.

(ii) $f_l < \frac{1}{2}$ refers to fragmentation, whereas $f_l > \frac{1}{2}$ is denoted as cratering.

(iii) Energy per volume S that yields $f_l = \frac{1}{2}$ is denoted as impact strength.

(iv) $f_{\rm KE} := 2E_{\rm kin}^{\rm frag}/E_{\rm kin}$: Fraction of the impact energy that is converted into kinetic energy of the fragments.

Different fragment sizes and velocities are summarised by appropriate distribution functions. m_i, D_i and v_i are mass, diameter and modulus of the velocity of a given fragment, respectively.

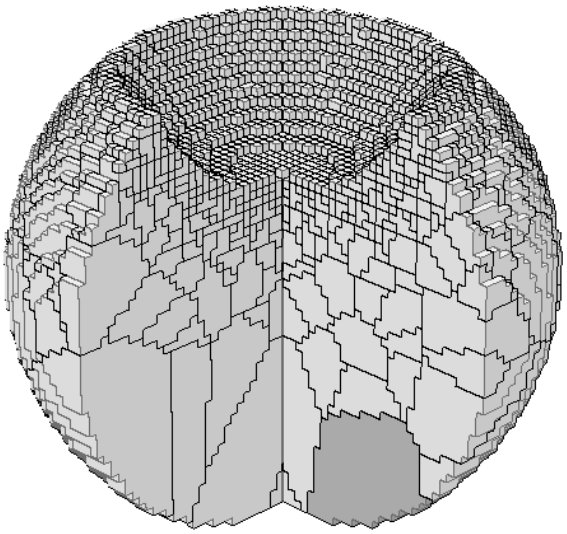

Figure 8. Section for $f_l = 0.04$ and $n = 3$. The largest fragment is coloured in dark-grey. In this calculation $60 \times 60 \times 60$ grid cells are used. Note the decomposition in grid cells and the Voronoy polyhedra which form the fragments.

(i) Fragment size distribution:

(a) $N_m(m)$: Number of all fragments with a mass $m_i \geqslant m$,
(b) $M(m)$: Mass of all fragments with a mass $m_i \geqslant m$,
(c) $N_D(D)$: Number of all fragments with a diameter $D_i \geqslant D$.

The distribution functions are related to each other:

$$\begin{aligned} N_m(m) &= N_D(D(m)) \\ M(m) &= \int_m^\infty x \left| \frac{\mathrm{d}N_m(x)}{\mathrm{d}x} \right| dx \\ N_m(m) &= \int_m^\infty \frac{1}{x} \left| \frac{\mathrm{d}M(x)}{\mathrm{d}x} \right| dx \end{aligned}$$

$D(m)$ is the size–mass relation.
(ii) Velocity distribution:

(a) $\bar{v}(m)$: mean velocity as a function of mass.

9.2 Prediction of collisional outcomes: Derivation from a Voronoy tessellation

Any theoretical or empirical prescription of a collision has to relate the afore mentioned parameters, namely the impact energy, to the sizes and velocities of the produced fragments. The central quantity is the impact strength, which is a measure for the overall stability of a body. Objects smaller than 1 m are accessible to laboratory experiments, while collisions of larger bodies up to asteroid size have to be analysed by complex computer simulations. Asteroid families, which are remnants of giant collisions in the asteroid belt, provide independent insight, although the data is difficult to interpret.

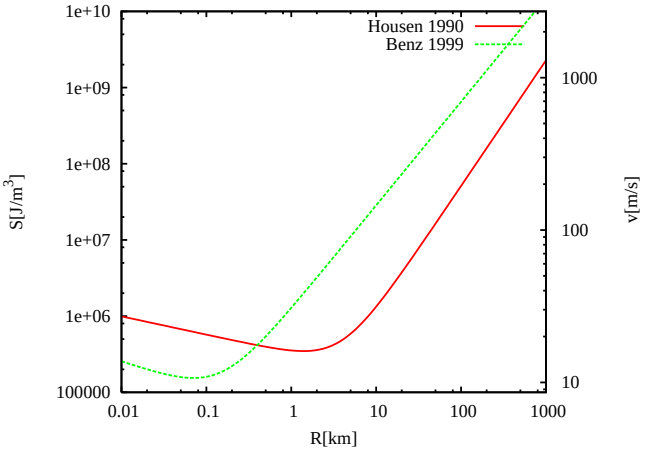

Figure 9. Impact strength according to Eq. 128 and Eq. 130. The right axis gives the corresponding impact velocity according to $S = 1/2\rho v^2$ with $\rho = 2.7$ g/cm^3.

We selected two different impact strength models as reference for our work. The first was obtained by Housen (1990) through the combination of asteroid family data and laboratory experiments via scaling laws:

$$S = S_0 \left(\frac{R}{1\,\mathrm{m}} \right)^{-0.24} \left[1 + 1.6612 \times 10^{-7} \left(\frac{R}{1\,\mathrm{m}} \right)^{1.89} \right]$$

$$f_{\mathrm{KE}} = 0.1 \tag{128}$$

$$S_0 = 1.726 \times 10^6\ \mathrm{Jm}^{-3} = 1.726 \times 10^7 \mathrm{erg\ cm}^{-3} \tag{129}$$

Later, Benz (1999) obtained another result through SPH simulations (for basalt, v =3 km/s):

$$S = S_0 \left(\frac{R}{1\,\mathrm{m}} \right)^{-0.38} \left[1 + 6.989 \times 10^{-5} \left(\frac{R}{1\,\mathrm{m}} \right)^{1.74} \right]$$

$$f_{\mathrm{KE}} \approx 0.01 \tag{130}$$

$$S_0 = 6.082 \times 10^5\ \mathrm{Jm}^{-3} \tag{131}$$

$$\rho = 2.7 \frac{\mathrm{g}}{\mathrm{cm}^3} \tag{132}$$

f_{KE} is a measure of the kinetic energy that is transferred to the fragments:

$$E_{\mathrm{kin}}^{\mathrm{frag}} = \frac{f_{\mathrm{KE}}}{2} E_{\mathrm{kin}} \tag{133}$$

We introduce a dimensionless measure γ of the relative importance of gravity for the result of a collision. It is defined as the ratio of the energy per volume S_G that is necessary to disperse the fragments to the impact strength S_0:

$$S_G = 2\pi \frac{4 - 2\sqrt[3]{2}}{5 f_{\mathrm{KE}}} G R^2 \rho^2$$

$$\gamma := S_G / S \tag{134}$$

The first step towards the prediction of a collisional outcome is to relate the impact energy and the impact strength to ascertain the size of the largest fragment f_l. Laboratory

experiments and simulations indicate the functional form

$$\epsilon(f_l) = \begin{cases} 2(1-f_l) & \text{for} \quad f_l > \frac{1}{2} \\ (2f_l)^{-\frac{1}{K}} & \text{otherwise.} \end{cases}$$
$$\epsilon = \frac{E_{\rm kin}\rho}{2SM} \tag{135}$$

which is both valid in the fragmentation regime and the cratering limit. The size of the largest fragment is used to derive the full size distribution. To accomplish the decomposition "seed fragments" are distributed inside the target according to the largest desired fragment. The full set of fragment is derived from a Voronoy tessellation [15] using these seed points. Fig. 8 depicts the result of such a decomposition. The fragment velocities are calculated from the total kinetic energy after the collision to initiate a post-collisional N–body calculation to treat reaccumulation.

We conducted a large set of such calculations to cover a sufficient range in f_l^i (i. e. impact energy) and γ (i. e. body size). Table 7 summarises the derived values of the largest and second largest fragment including reaccumulation.

10 COLLISIONAL CASCADES

A first well-defined application of the fragmentation model is a collisional cascade. The term *cascade* denotes that fragments of one collision in a many-body system may hit other bodies, whose fragments further shatter even more bodies. Thus the particle number increases exponentially with every subsequent collision.

Although the formation of planets requires a net growth due to collisions, this destructive process plays a role in the formation of larger bodies as the overall size distribution controls the accretion rate of the protoplanets. Therefore it is worth to have a closer look into this mechanism.

10.1 Self–similar collisions

A system of colliding bodies is usually embedded in a broader context, like stars moving in a galaxy or asteroids orbiting in our own solar system. First, we simplify this dynamical background as well as some aspects of the collisions to make the problem tractable.

The first step is to decompose an inhomogeneous system into smaller subvolumes which are locally homogenous. Furthermore, it is assumed that these subvolumes hardly interact with each other. Hence it is possible to apply the *particle–in–a–box–method* (Safronov 1969) to analyse collisions within the small subvolumes:

(i) All particles are contained in a constant volume.
(ii) The particle sizes are described by a distribution function $n(m)$, i. e. the particle number per volume and mass interval.
(iii) For convenience, we assume a constant (or typical) relative velocity for a given pair of colliding bodies.

[15] The Voronoy tessellation assigns every volume element to the closest seed point. First applications date back to the 17th century, but the Russian mathematician of Ukrainian origin Georgy Feodosevich Voronoy put it on a general base in 1908.

The distribution function is evolved by the coagulation equation. We modified the equation given by Tanaka et al. (1996) by introducing a new function $M_{\rm red}$ to arrive at a more concise expression:

$$0 = \frac{\partial}{\partial t} m n(t,m) + \frac{\partial}{\partial m} F_m(t,m) \tag{136}$$

The mass flux F_m is given by:

$$F_m = -\iint n(t,m_1) n(t,m_2) \xi dm_1 dm_2$$
$$\xi \equiv \sigma(m_1,m_2) v_{\rm rel} M_{\rm red}(m,m_1,m_2) \tag{137}$$
$$M_{\rm tot} = \int n(t,m) m dm \tag{138}$$
$$\frac{\partial}{\partial t} M_{\rm tot} = F_m(m_{\rm min}), \tag{139}$$

where ξ is the coagulation kernel, n is the already introduced size distribution,$v_{\rm rel}$ is the mean relative velocity, σ is the cross section for colliding bodies (m_1, m_2) and $M_{\rm red}$ is the newly introduced fragment redistribution function. $M_{\rm red}$ contains all information on the fragments arising from the breakup of body m_1 due to the impact of body m_2. Its definition avoids double counting of collisions in the above integral. The redistribution function is related to the differential number distribution function $n_{\rm coll}(m_1,m_2,m)$, i. e. the number of fragments produced by a collision per mass interval. Since the target m_1 formally disappears, it is included as a negative contribution:

$$M_{\rm red}(m,m_1,m_2) := \int_0^m \big(n_{\rm coll}(m_1,m_2,\tilde{m}) - \delta(\tilde{m}-m_1)\big)\tilde{m} d\tilde{m} \tag{140}$$

Mass conservation in each collision is reflected by $M_{\rm red}(0,m_1,m_2) = M_{\rm red}(\infty,m_1,m_2) = 0$. The cross section σ depends on the velocities and radii R_i of the particles. A simple approach is the geometric cross section:

$$\sigma(m_1,m_2) = \pi(R_1+R_2)^2 \tag{141}$$

If gravity plays an important role during encounters, two colliding bodies move on hyperbolic orbits with a pericentre distance that is smaller than the impact parameter. This leads to an additional enlargement of the cross section, denoted as gravitational focusing:

$$\sigma(m_1,m_2) = \pi(R_1+R_2)^2 \left(1 + \frac{2G(m_1+m_2)}{v_{\rm rel}^2 (R_1+R_2)}\right) \tag{142}$$

A special class of collisional models are self-similar collisions. Self-similarity implies an invariance of the collisional outcome with respect to the scale of the colliding bodies. If the target mass as well as the projectile mass are enlarged by a factor of two, then only the masses of all fragments doubles without further changes in the collisional outcome. They allow us to introduce the convenient dimensionless fragment redistribution function f_m:

$$M_{\rm red}(m,m_1,m_2) = m f_m(m_1/m, m_2/m) \tag{143}$$

We follow Tanaka et al. (1996) and employ the substitution[16] $m_1 = m x_1$, $m_2 = m x_2$ to simplify Eq. 137:

[16] A similar approach to the solution of the coagulation equation is the Zakharov transformation, see Connaughton et al. (2004).

Largest Fragment

γ f_l^i	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
0.02552	0.10000	0.20000	0.30000	0.40000	0.50000	0.60000	0.70000	0.80000
0.19897	0.10000	0.20000	0.30000	0.40000	0.50000	0.60000	0.70000	0.80000
0.67985	0.10000	0.20000	0.30000	0.40000	0.57612	0.67315	0.82006	0.96073
1.14050	0.10000	0.31526	0.35708	0.61362	0.83511	0.92832	0.94380	0.97572
1.77057	0.10884	0.58883	0.75974	0.87922	0.92755	0.93662	0.97107	0.97924
2.26021	0.15895	0.68891	0.87217	0.89592	0.92965	0.96089	0.96701	0.98727
3.11626	0.30954	0.83774	0.90272	0.92682	0.95943	0.95565	0.97677	0.98791

Second largest Fragment

γ f_l^i	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8
0.02552	0.08171	0.10770	0.06471	0.08107	0.04976	0.03982	0.03171	0.02155
0.19897	0.09174	0.09533	0.08410	0.06967	0.04930	0.04825	0.03510	0.02077
0.67985	0.07713	0.08365	0.07791	0.08387	0.07026	0.06147	0.06082	0.00847
1.14050	0.08621	0.07256	0.10331	0.09640	0.02467	0.00675	0.00783	0.00265
1.77057	0.07909	0.05549	0.06961	0.02035	0.00329	0.00719	0.00273	0.00161
2.26021	0.06693	0.02288	0.00528	0.00882	0.00584	0.00268	0.00664	0.00126
3.11626	0.06940	0.00884	0.00384	0.00488	0.00064	0.01007	0.00225	0.00162

Table 7. Data compilation of the fragmentation calculations.

$$F_m = -\iint n(t, mx_1)n(t, mx_2)m^{11/3} \sigma(x_1, x_2)v_{\rm rel}f_m(x_1, x_2)dx_1dx_2 \quad (144)$$

A simple solution is a steady-state cascade with $F_m = \text{const}$. The loss of bodies of a given size is balanced by the fragment supply from larger bodies, hence the system maintains a steady-state $\frac{\partial}{\partial t}n(t, m) = 0$. Eq. 144 inspires the ansatz $n(m) \propto m^{-k}$, which yields $k = 11/6$. This is the well known equilibrium slope in self-similar collisional cascades, which was already found by Dohnanyi (1969). Strong gravitational focusing changes the exponent[17] to $k = 13/6$. Both steady-state solutions seem to be rather artificial, as they contain an infinite amount of mass and require a steady mass influx from infinity. However, they provide an appropriate description for the relaxed fragment tail of a size distribution, as long as the largest bodies provide a sufficient flux of new fragments. Once the largest bodies start to decay, the finite amount of mass in the system leads to an overall decay of the collisional cascade. Thus we seek for a more general solution to Eq. 136 using the ansatz $n(t, m) = a(t)n_0(m)$:

$$\frac{\partial}{\partial t}a(t) = -Ca(t)^2 \quad (145)$$

$$mn_0(m) = \frac{1}{C}\frac{\partial}{\partial m}F_m \quad (146)$$

C is determined by fixing n_0 at an arbitrary value m^*. $a(t)$ is independent of the collision model:

$$a(t) = \frac{1}{1 + Ct} \quad (147)$$

$$C \propto n(m^*) \quad (148)$$

A power law solution is $n_0(m) \propto -Cm^{-k+1}$ which is only valid for $C < 0$ (agglomeration dominates). To examine $C > 0$, we perturb the already known equilibrium solution:

$$n_0(m) = N_0m^{-k} - CN_1m^{-2k+2} + \mathcal{O}(C^2) \quad (149)$$

$$\frac{1}{N_1} = (2 - k)\iint x_1^{-k}x_2^{-2k+2}\sigma(x_1, x_2) v_{\rm rel}\left(f_m(x_1, x_2) + f_m(x_2, x_1)\right)dx_1dx_2 \quad (150)$$

N_1 is small if the integral on the right hand side is large. This is the case for a sufficiently large impact strength. Eq. 149 has the interesting property that $n(m') = 0$ for some mass m', given that $k < 2$. This mass m' represents the largest body in the system, e. g. the largest asteroid in a fictitious asteroid belt.

[17] Tanaka et al. (1996) state that $k < 2$ is a necessary condition for a finite mass flux. However, their analysis is not valid for all possible collisional models.

11 SIZE–DEPENDENT STRENGTH

Self-similarity is an enormous help in analysing the coagulation equation. It releases completely the need to know any specific details of the collisional process and provides valuable insight at the same time. But self-similarity is also a strong limitation on the underlying collisional physics.

A major component of a fragmentation model is the knowledge of the impact strength as a function of size. Simulations as well as asteroid families establish that it is not some fixed value, but changes with size which immediately breaks the self-similarity. Larger bodies are weaker due to an increasing number of flaws (there are no big monocrystals), but then gravity leads to a turnover and increases the strength.

We model the size dependent strength S with a power law to examine the influence on the equilibrium solution. The velocity dispersion v and the collisional cross section σ are also modelled with power laws to account for relaxation processes:

$$v = v_0 \left(\frac{m}{m_0}\right)^w \tag{151}$$

$$\sigma = \sigma_0 \left(\frac{m}{m_0}\right)^s \tag{152}$$

$$S = S_0 \left(\frac{m}{m_0}\right)^\alpha \tag{153}$$

The subscript "0" denotes values for an arbitrarily chosen scaling mass. Since smaller bodies are more abundant than larger ones, we safely assume that most collisions involve a large mass ratio. In addition, we assume $w < 0$, since we expect energy equipartition to some degree in most cases. These restrictions lead to the following simplifications ($m_1 > m_2$):

$$\sigma(m_1, m_2) \approx \sigma(m_1) \tag{154}$$

$$v_{\rm rel} \approx v(m_2) \tag{155}$$

$$\epsilon \approx \frac{1}{2}\frac{m_2 \rho v_{\rm rel}^2}{m_1 S_1} \tag{156}$$

Therefore the smaller body m_2 enters only through the specific energy ϵ:

$$F_m \approx -\iint n(t,m_1) n(t,m_2) \sigma(m_1) v_{\rm rel}(m_2) m_1 \\ f_m(m_1/m, \epsilon) dm_1 dm_2 \tag{157}$$

We introduce new dimensionless quantities with the help of Eq. 154–156 to simplify the integral:

$$m_1 = m x_1$$

$$m_2 = m_0 \left(\frac{m_1}{m_0}\right)^{\frac{1+\alpha}{1+2w}} \left(\frac{2S_0}{\rho v_0^2}\right)^{\frac{1}{1+2w}} \epsilon^{\frac{1}{1+2w}} \tag{158}$$

Again we assume a power law for the density $n \propto m^{-k}$ and change the integration parameters to (x_1, ϵ). Applying the constant–flux condition yields the equilibrium exponent

$$k \approx \frac{s + 3 + \alpha + w(2s + \alpha + 5)}{2 + \alpha + 2w} \tag{159}$$

and the scaling exponent k' of the total mass loss:

$$k' \approx \frac{s - w + 1}{2 + \alpha + 2w}$$

$$\frac{\partial}{\partial t} M_{\rm tot} \propto -\tilde{S}^{-k'} \qquad \tilde{S} = \frac{2S_0}{\rho v_0^2} \tag{160}$$

The exponent k' in Eq. 160 is close to unity for realistic values of the free parameters. Thus the mass loss is roughly inversely proportional to the strength of the bodies. The general formula Eq. 159 contains the special solution of O'Brien (2003), who concentrated on the parameters $s = 2/3$, $w = 0$ and a special collisional model. In fact, the derivation applies to a much wider class of collisional models that we denote as scalable collisional models. Scalable indicates that the model is self-similar except a scaling of the impactor mass.

12 PERTURBATION OF EQUILIBRIUM

The derived scaling relations provide insight into the overall properties of a collisional cascade, which is in (or close to) equilibrium. However, they do not provide information on how the equilibrium is attained or how the system responds to various external perturbations. A rigorous approach would be the approximate solution of the coagulation equation[18], which is by no means simple since it requires a careful analysis of the collision model.

Hence we turn to perturbations of the equilibrium size distribution, as it is easier to asses the quality of the derived expression for a variety of collision models. In addition, all equations are linear in the perturbation, allowing the detailed analysis of the solution.

If the equilibrium solution $n(m) = n_0 (m/m_0)^{-k}$ is perturbed with a small deviation $\Delta n(m)$, we get to first order:

$$0 = \frac{\partial}{\partial t} m \Delta n(m) + \frac{\partial}{\partial m} F_p(t, m) \tag{161}$$

$$F_p = -\iint \Delta n(m_1) n(t, m_2) \sigma(m_1, m_2) v_{\rm rel} \\ \times \left(M_{\rm red}(m, m_1, m_2) + M_{\rm red}(m, m_2, m_1)\right) dm_1 dm_2$$

Despite of the expansion in Δn, Eq. 161 is still a complicated integro-differential equation. Thus it is not possible to obtain a solution without further information about the problem. While there is no general solution, we restrict our attention to self-similar collisional processes. In virtue of this assumption it is possible to simplify Eq. 161, as we can see in Eq. 162 and 162. In those expressions σ_0 and v_0 are velocity and cross section of an arbitrarily chosen scaling mass m_0. $F(x_1)$ contains all information about the collisional process. If collisions do not result in extreme outcomes, like cratering or a complete destruction of the target, most of the fragment mass is contained in bodies with similar size as the parent body. Hence we expect that $F(x_1)$ peaks around $x_1 \approx 1$ and drops to zero as x_1 gets larger (or smaller). We introduce the dimensionless relative perturbation $g(m)$:

$$g(m) = \frac{\Delta n(m)}{n(m)} = \frac{\Delta n(m) m^k}{n_0 m_0^k} \tag{164}$$

[18] Appendix C highlights a possible approach.

$$0 = \frac{\partial}{\partial t} m\Delta n(m) - n_0 m_0^3 \sigma_0 v_0 \frac{\partial}{\partial m} \int \Delta n(t, mx_1) F(x_1) (mx_1/m_0)^k dx_1 \tag{162}$$

$$F(x_1) = \int m_0^{2k-3} x_1^{-k} x_2^{-k} \frac{\sigma(x_1, x_2)}{\sigma_0} \frac{v_{\rm rel}}{v_0} (f_m(x_1, x_2) + f_m(x_2, x_1)) dx_2 \tag{163}$$

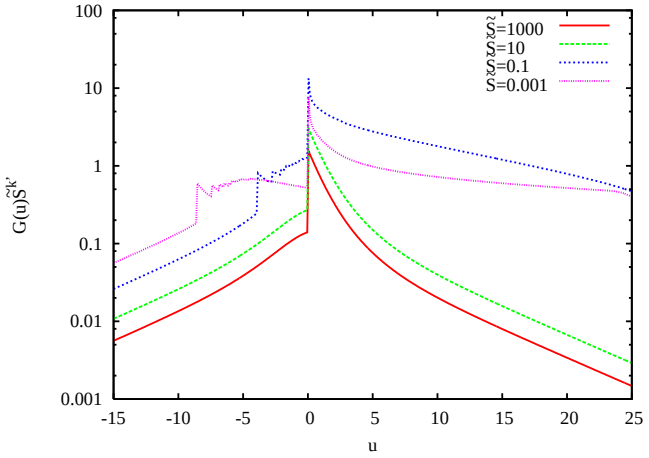

Figure 10. Scaled fragmentation kernel $G(u)$ for a simple fragmentation model (see Eq. 200) and different scaled impact strength $\tilde{S}$.

Thus the new differential equation reads:

$$0 = \frac{\partial}{\partial t} (m/m_0)^{1-k} \Delta g(m) - n_0 m_0^2 \sigma_0 v_0 \frac{\partial}{\partial m} \int \Delta g(t, mx_1) F(x_1) dx_1 \tag{165}$$

We change to logarithmic coordinates to arrive at a convolution integral:

$$u = \ln(m/m_0) \qquad u_1 = \ln(x_1) \tag{166}$$

Furthermore we define a collisional timescale τ_0

$$\tau_0 = (n_0 m_0 \sigma_0 v_0)^{-1} \tag{167}$$

to obtain a more concise expression. The transformed equation is:

$$0 = \frac{\partial}{\partial t} g(t, u) e^{u(2-k)} - \frac{1}{\tau_0} \frac{\partial}{\partial u} \int g(t, u + u_1) G(u_1) du_1 \tag{168}$$

$$G(u) = F(e^u) e^u \tag{169}$$

If $g(u)$ is varying on a scale larger than the width of the kernel $G(u)$ (compare Fig. 10), it is justified to expand $g(u)$ under the integral. We retain the first two moments of $G(u)$:

$$0 = \frac{\partial}{\partial t} g(t, u) e^{u(2-k)} - \frac{G_0}{\tau_0} \frac{\partial}{\partial u} g(t, u) - \frac{G_1}{\tau_0} \frac{\partial^2}{\partial u^2} g(t, u) \tag{170}$$

$$G_k = \int u^k G(u) du \tag{171}$$

The first order moment G_1, which introduces a diffusive term, is omitted in the following for clarity[19]. We introduce a fragmentation time $\tau_{\rm frag}(u)$ and transform Eq. 170 back to m:

$$0 = \frac{\partial}{\partial t} g(t, m) - \frac{m}{\tau_{\rm frag}(m)} \frac{\partial}{\partial m} g(t, m) \tag{172}$$

$$\tau_{\rm frag} = \frac{\tau_0}{G_0} e^{u(2-k)} = \frac{\tau_0}{G_0} (m/m_0)^{2-k} \tag{173}$$

Eq. 172 is a modified advection equation, which conserves the total mass. It is possible to derive equations similar to Eq. 172 for any collisional model. However, the general approach is less fruitful, as it lacks a robust frame of a known equilibrium solution and reliable scaling relations. Therefore we provide only the extension to scalable collisional models in Appendix B. We readily obtain the general solution:

$$g(t, m) = f\left(t + \tau_0 \frac{(m/m_0)^{(2-k)}}{G_0(2-k)}\right)$$

$$\Delta n(t, m) = n(m) f\left(t + \tau_0 \frac{(m/m_0)^{(2-k)}}{G_0(2-k)}\right) \tag{174}$$

The function f is determined by the initial value $g(0, m)$ of the perturbation. As the collisional cascade evolves, the initial perturbation function is shifted as a whole to smaller masses. This evolution becomes clearer if we attach labels $M(0)$ to the initial perturbation function and follow the time evolution of these tags. The functions $M(t)$ are the characteristics[20] of the differential equation 172:

$$M(t) = m_0 \left[(M(0)/m_0)^{(2-k)} - t/\tau_0 G_0 (2-k)\right]^{1/(2-k)} \tag{175}$$

The meaning of the fragmentation time $\tau_{\rm frag}$ becomes clear by the relation

$$\frac{M}{\dot{M}} = -\tau_{\rm frag} \tag{176}$$

which is the time until a body has lost a significant fraction of its mass due to destructive collisions. A comparison of the perturbation equation 172 with the scaling relations from the previous section gives the scaling of the zeroth order moment G_0 with respect to the impact strength:

$$G_0 = G_0' \tilde{S}^{-k'} \tag{177}$$

G_0' should only depend on the fragmentation model (i. e. fragment size distribution as a function of the largest fragment f_l) within the limits of this approximation. Fig. 10 shows that the scaling with the impact strength works quite well, except slight variations which are small compared to the covered range of impact strengths. Likewise, it is possible to restate the total equilibrium flux $F_{\rm eq}$ in terms of G_0':

$$F_{\rm eq}(m) \approx -\frac{G_0'}{2} n(m)^2 \sigma(m) m^3 v_{\rm rel} \tilde{S}^{-k'} \tag{178}$$

[19] The study of wave-like structures in the size distribution (Bagatin et al. 1994,see e.g.) requires even the second order moment G_2.

[20] In general, characteristics of a partial differential equation are paths along which the solution is constant.

The fragmentation timescale $\tau_{\rm frag}(m)$ allows a more intuitive expression:

$$F_{\rm eq}(m) \approx -\frac{1}{2}\frac{n(m)m^2}{\tau_{\rm frag}(m)} \tag{179}$$

Our simple collisional model (see Fig. 10 and Eq. 200) refers to:

$$F_{\rm eq}(m) = -(1\ldots30)\times n(m)^2\sigma(m)m^3 v_{\rm rel}\tilde{S}^{-k'} \tag{180}$$

13 MIGRATION AND COLLISIONS

The local perturbation analysis is only applicable to a planetesimal disc, if the migration velocity of the planetesimals is negligible small. This assures that collisional cascades at different radial distances do not couple to each other, so that the whole disc is composed of many local cascades. While this assumption is justified for larger bodies, migration is strongly influencing bodies below 1 km in size. Hence we extend our analysis to examine the influence of migration on the (no longer) local collisional processes.

We assume that the collisional evolution of the system leads to an equilibrium planetesimal distribution everywhere in the disc:

$$\Sigma_0(r,m) = \Sigma_{r,0}(r)C_0(m) \tag{181}$$

$\Sigma_r(r)$ is the total surface density at a given distance r, while $C_0(m)$ is the universal equilibrium distribution. Though the planetesimal distribution at larger sizes is likely different at different locations in the disc, we only demand a universal function at smaller sizes, where migration is important. The power law exponent k depends on the details of the invoked physics, but numerical simulations show that $k \approx 2$ is a fiducial value. Eq. 181 does not yet include migration effects. If we include migration, the surface density is modified to

$$\Sigma(r,m) = g(r,m)\Sigma_0(r,m) \tag{182}$$

where the dimensionless function g contains the changes due to migration. The collisional evolution is governed by the continuity equation with an additional collisional term

$$\frac{\partial\Sigma(r,m)}{\partial t} - \frac{1}{r}\frac{\partial}{\partial r}\left(v(r,m)r\Sigma(r,m)\right) = \dot{\Sigma}_{\rm coll} \tag{183}$$

where $v(r,m)$ is the migration velocity (see Eq. 77), defined such that positive v imply an inward migration. We express the collisional term with the help of Eq. 172 and seek for a steady-state solution $\dot{\Sigma} = 0$:

$$\frac{1}{\tau_{\rm frag}(m,r)}\frac{\partial g}{\partial m}m\Sigma_{r,0}(r) + \frac{1}{r}\frac{\partial}{\partial r}\left(gvr\Sigma_{r,0}(r)\right) = 0 \tag{184}$$

$\tau_{\rm frag}(m,r)$ is the fragmentation timescale of a mass m at a distance r. Since the surface density Σ and the various contributions to the drag force are well described by a power law (with respect to radius), Eq. 184 further simplifies to:

$$\frac{1}{\tau_{\rm frag}(m,r)}\frac{\partial g}{\partial m}m + \frac{\partial g}{\partial r}v - \frac{b}{r}gv = 0 \tag{185}$$

b is a combination of the various invoked power law exponents. As the surface density Σ and the gas density drop with increasing radius in any realistic disc model, it is safe to assume $b > 0$. We choose a self-similar ansatz for g:

$$g(r,m) = g(\zeta) \quad , \quad \zeta = mg_m(r) \tag{186}$$

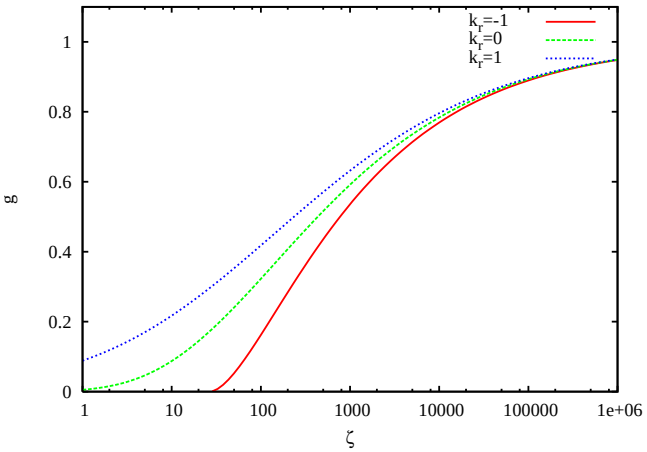

Figure 11. Cut-off function g according to Eq. 192. The mass exponent is $k_m = 1/3$, while the mass influx exponent is $b = 1.75$ according to the minimum solar nebula.

The new differential equation is

$$\frac{1}{\tau_{\rm frag}(m,r)}\frac{dg}{d\zeta}mg_m(r) + m\frac{dg}{d\zeta}\frac{dg_m}{dr}v - \frac{b}{r}gv = 0 \tag{187}$$

which is equivalent to the more concise expression:

$$\frac{d\ln(g)}{d\ln(\zeta)}\left(\frac{r}{v\tau_{\rm frag}} + \frac{d\ln(g_m)}{d\ln(r)}\right) = b \tag{188}$$

We assume a power-law dependence for the timescale ratio $\tau_{\rm mig}/\tau_{\rm frag}$:

$$\frac{r}{v\tau_{\rm frag}} = \frac{\tau_{\rm mig}}{\tau_{\rm frag}} = (m/m_0)^{k_m}(r/r_0)^{k_r} \tag{189}$$

The cut-off mass m_0 at a distance r_0 has a timescale ratio $\tau_{\rm mig}/\tau_{\rm frag} = 1$, which defines a proper lower cut-off within this context. Hence the solution is:

$$b = \frac{d\ln(g)}{d\ln(\zeta)}\left(\zeta^{k_m} + \frac{k_r}{k_m}\right) \tag{190}$$

$$g_m(r) = \frac{(r/r_0)^{k_r/k_m}}{m_0} \tag{191}$$

$$g(\zeta) = \left(1 + \frac{k_r}{k_m\zeta^{k_m}}\right)^{-b/k_r} \tag{192}$$

Though the analytical solution Eq. 192 provides a complete description of the lower cut-off of the size distribution, it is more appropriate within the frame of this discussion to translate the equilibrium solution to an equilibrium mass loss due to migration:

$$\dot{\Sigma}_{\rm mig}(r,m) = -\frac{bv}{r}\Sigma + \frac{bv}{r}\Sigma\frac{k_r/k_m}{\zeta^{k_m} + k_r/k_m} = -\frac{b\Sigma}{\tau_{\rm mig} + k_r/k_m\tau_{\rm frag}} \tag{193}$$

An inspection of the timescale ratio shows that the mass exponent k_m should be positive, whereas simple estimations of k_r on the basis of the minimum mass solar nebula are somewhat inconclusive. The value of k_r is so close to zero that any change in the assumed equilibrium slope or the impact strength scaling gives easily both positive and negative values. Moreover, Eq. 193 requires a globally relaxed planetesimal disc, but the huge spread in the various involved timescales at different radii inhibits any significant relaxation in the early stages.

However, it is possible to gain valuable information from the two limiting cases $k_r > 0$ and $k_r < 0$. Both values of k_r give the proper limit $g \to 1$ at large masses, where the migration timescale is much larger than the fragmentation timescale and we recover the steady-state collisional cascade.

A positive exponent k_r reduces the effective mass loss due to migration, as fragments from the outer part of the disc replenish the local mass loss. Hence the fragmentation timescale controls the net loss of smaller planetesimals. In contrast, a negative exponent k_r leads to a pronounced cut-off in the size distribution, since only larger planetesimals are replenished through inward migration. Though the mass loss rate is singular at some mass m', this sharp cut-off is an artifact due to the perturbation approximation.

Our analysis is subjected to several restrictions. We applied the perturbation equation to values of g that exceed the limit for a safe application (i. e. $g \not\approx 1$) of the perturbation expansion. Furthermore, the steady-state solution requires a global relaxation of the collisional processes, which is practically never obtained during the disc evolution. Despite of these restrictions, we gained insight on a more qualitative level. Numerical calculations indicate that the perturbation approximation is inappropriate close to the lower cut-off of the size distribution. However, a comparison of different exponents k_r (see Fig. 11) attributes only a minor role to the replenishment of fragments due to inward migration. Only unrealistic small slopes b of the migrational mass influx would strengthen the importance of this process. Though temporally non-equilibrium phenomena are not ruled out by the previous derivation, their study would require the global simulation of the system.

14 COAGULATION

While most coagulation kernels are only restricted to a limited analytical analysis (e. g. scaling relations), there exist some special kernels that allow the closed solution of the coagulation equation. All rely on the assumption of perfect mergers, which allows the reformulation of the general equation 136 to

$$\frac{\partial n(m,t)}{\partial t} = \frac{1}{2}\int_0^m \mathcal{K}(m-m',m')n(m-m',t)n(m',t)dm' - n(m,t)\int_0^\infty \mathcal{K}(m,m')n(m',t)dm'$$

where $\mathcal{K}$ is the coagulation kernel. One of these particular kernels was introduced by Safronov (1969):

$$\mathcal{K}(m_1,m_2) = A_1(m_1+m_2) \tag{194}$$

This coagulation kernel implies perfect mergers, where the coalescence rate of two particles m_1 and m_2 is assumed to be proportional to the sum of their masses. It seems that this is an artificial choice, devised to allow an analytic solution. However, the Safronov cross section provides an intermediate case between a geometric cross section ($\sigma \propto m^{2/3}$) and strong gravitational focusing ($\sigma \propto m^{4/3}$). A special solution to the initial condition

$$n(m,0) = \frac{n_0}{m_0}\exp\left(-m/m_0\right) \tag{195}$$

ρ	2,700 kg/m^3
k	1/6
Model	Gaussian Scatter
f_{KE}	0.1
K	1.24

Table 8. Main parameters of the collisional model.

is the function (Ohtsuki et al. 1990,see e.g.)

$$n(m,\tau) = \frac{n_0\tilde{g}}{m\sqrt{1-\tilde{g}}}\exp(-(2-\tilde{g})m/m_0)I_1(2m/m_0\sqrt{1-\tilde{g}}) \tag{196}$$

$$\tilde{g} = \exp(-\tau) \qquad \tau = A_1\rho t$$
$$\rho = \int_0^\infty mn(m)dm = n_0m_0 \tag{197}$$

where τ is the dimensionless time and I_1 is a modified Bessel function of the first kind.

15 MODELS FOR $M_{\rm RED}$

Though we already obtained insight into the nature of collisional cascades without a detailed specification of the coagulation kernel, any detailed study of a collisional system requires the specification of a realistic collisional model.

First, we restate the well-known perfect accretion model. While it is a gross oversimplification for collisions among kilometre–sized planetesimals, its simplicity allows a reliable code testing and eases the comparison with other works:

$$M_{\rm red}(m,m_1,m_2) = -m_1\Theta(m-m_1) - m_2\Theta(m-m_2) + (m_1+m_2)\Theta(m-m_1-m_2) \tag{198}$$

Although our fragmentation model (see section 9) provides a very detailed description of the outcome of a collision, we abandon most of the details for the following reasons. The computational effort of the numerical solution of the coagulation equation scales with the third power of the number of mass bins. Hence we chose a mass grid whose resolution is by far smaller than the information provided by the detailed collisional model. As a mismatch of the mass resolution could produce undesired artifacts, a lower resolution of the collisional model is needed for consistence. Thus only the largest fragment $f_l(f_l^i,\gamma)$ and the second fragment $f_l^{(2)}(f_l^i,\gamma)$(which contains information on reaccumulation) enter the fragment size distribution:

$$\frac{M_{\rm red}(xM)}{M} = \begin{cases} 1 & \text{if } x \geqslant f_l \\ 1-f_l & \text{if } f_l > x \geqslant f_l^{(2)} \\ (1-f_l-f_l^{(2)})(x/f_l^{(2)})^{f_l} & \text{otherwise} \end{cases} \tag{199}$$

Both values f_l and $f_l^{(2)}$ are interpolated from table 7, where the initial fragment size f_l^i is calculated from the dimensionless impact energy ϵ. We used a reduced fragmentation model for test purposes:

$$M_{\rm red}(xM)/M = \begin{cases} 1\text{if } x \geqslant f_l \\ (1-f_l)(x/f_l)^{f_l}\text{otherwise} \end{cases} \tag{200}$$

Table 8 summarises the most important model parameters.

16 STATISTICAL MODEL

The direct approach to the integration of an N–body system is, in principle, possible for any particle number. While this procedure becomes computationally too expensive for very large particle numbers, a by far more efficient approach is applicable in this regime. Instead of tracking all particle orbits, a distribution function f (also phase-space density), which gives the probability to find a particle at a position $\mathbf{x}$ with a velocity $\mathbf{v}$, contains the state of the system:

$$dp = f(\mathbf{x}, \mathbf{v}) d^3x d^3v \tag{201}$$

As long as only dynamical interactions are taken into account, the number of all particles (e. g. stars, planetesimals) is conserved. The continuity equation reads:

$$0 = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{v}} \tag{202}$$

This is the collisionless Boltzmann equation. Collisions lead to an additional term

$$\left(\frac{\partial f}{\partial t}\right)_{\rm coll} = \frac{\partial f}{\partial t} + \mathbf{v} \cdot \nabla f - \nabla \Phi \cdot \frac{\partial f}{\partial \mathbf{v}} \tag{203}$$

which will be discussed later. f is a function of six variables, so an exact solution is usually very complicated or even impossible. However, it is possible to gain valuable insight into the problem by taking the moments of the distribution function:

$$\langle x_i^n v_j^m \rangle = \int f(\mathbf{x}, \mathbf{v}) x_i^n v_j^m d^3x d^3v \qquad n, m > 0 \tag{204}$$

The spatial density (particles per volume) is related to distribution function:

$$\nu(\mathbf{x}) = \int f(\mathbf{x}, \mathbf{v}) d^3v \tag{205}$$

Integration of Eq. 203 over all velocities yields the corresponding continuity equation:

$$\frac{\partial \nu}{\partial t} + \frac{\partial \nu \bar{v}_i}{\partial x_i} = \left(\frac{\partial \nu}{\partial t}\right)_{\rm coll} \tag{206}$$

The first order moment with respect to velocity gives the time evolution of the mean velocity $\bar{v}$

$$\nu \frac{\partial \bar{v}_j}{\partial t} + \nu \bar{v}_i \frac{\partial \bar{v}_j}{\partial x_i} = -\nu \frac{\partial \Phi}{\partial x_j} - \frac{\partial (\nu \sigma_{ij}^2)}{\partial x_i} + \nu \left(\frac{\partial \bar{v}_j}{\partial t}\right)_{\rm coll} \tag{207}$$

$$\bar{v}_i = \frac{1}{\nu} \int f(\mathbf{x}, \mathbf{v}) v_i d^3v$$

$$\sigma_{ij}^2 = \overline{v_i v_j} - \bar{v}_i \bar{v}_j \tag{208}$$

where σ_{ij} is the anisotropic velocity dispersion and the continuity equation was used to arrive at a more concise formulation. Equations 206 and 207 are the Jeans equations. While the structure of the moment equations is already familiar from hydrodynamics, they do not provide a closed set of differential equations, since each differential equation of a given moment is related to (yet unknown) higher order moments. Hence any finite set of momenta needs a closure relation – additional constraints that relate the highest order moments to known quantities. The choice of this relation is a key element in the validity of the equations, but it is not unique and depends well on the problem at hand (Larson 1970,compare e.g.).

Owing to the geometry of a planetesimal disc, it is useful to express the Boltzmann equation in cylindrical coordinates

$$\frac{\partial f}{\partial t} + v_r \frac{\partial f}{\partial r} + v_z \frac{\partial f}{\partial z} + \left(\frac{v_\phi^2}{r} - \frac{\partial \Phi}{\partial r}\right) \frac{\partial f}{\partial v_r} - \frac{v_r v_\phi}{r} \frac{\partial f}{\partial v_\phi} - \frac{\partial \Phi}{\partial z} \frac{\partial f}{\partial v_z} = 0 \tag{209}$$

where all derivatives with respect to ϕ have been dropped due to the assumed axisymmetry of the disc.

16.1 Distribution Function

Any statistical description of a planetesimal disc requires the knowledge of the distribution function. Since the full problem including collisions, encounters and gas drag has no analytic solution, a collisionless planetesimal disc (i. e. no perturbations) is a natural basis for further investigations. The distribution function that describes such a simplified system is a solution of the Boltzmann equation. A special solution to Eq. 209 is a thin homogenous planetesimal disc

$$f(z, v) = \frac{\Omega \Sigma}{2\pi^2 T_r T_z m} \exp\left(-\frac{v_r^2 + 4v_\phi^2}{2T_r} - \frac{v_z^2 + \Omega^2 z^2}{2T_z}\right) \tag{210}$$

provided that the radial velocity dispersion T_r and the vertical dispersion T_z are small compared the mean orbital velocity v_K. The azimuthal velocity dispersion T_ϕ is locked to T_r by the local epicyclic frequency κ in a central potential, where the ratio $1 : 4$ is a special solution of (Binney 1994,see e.g)

$$\kappa^2 T_r = 4\Omega^2 T_\phi \tag{211}$$

All velocities v_r, v_ϕ and v_z refer to the local Keplerian velocity. The normalisation is the same as in Stewart (2000):

$$\int d^3v dz f(z, v) = \frac{\Sigma}{m} \tag{212}$$

A planetesimal disc is a slowly evolving system compared to the orbital time, hence it is reasonable to use Eq. 210 as a general solution of the perturbed problem. Σ, T_z and T_r are now functions of time and of the radial distance to the star. All information on the system is contained in these three momenta of the distribution function, where higher order moments can be deduced from Eq. 210. Thus the functional form of the distribution function represents an implicit closure relation.

The validity of this approximation can be further assessed by a closer examination of the Boltzmann equation. We summarise all perturbations in an evolution timescale $T_{\rm evol}$ and reduce the radial structure to some typical length scale Δr to estimate the deviation from the functional form Eq. 210. A comparison with Eq. 209 shows that the difference is small if the migration timescale and the evolution timescale are large compared to the orbital time T_0:

$$T_0 \ll \Delta r / \langle v_r \rangle \tag{213}$$

$$T_0 \ll T_{\rm evol} \tag{214}$$

An order–of–magnitude estimate of the evolution time supports condition 213 and 214. Furthermore, numerical calculations confirm that the velocity distribution stays triaxial Gaussian (see Ida 1992).

The distribution function is equivalent to an isothermal vertical density structure with scale height h:

$$h = \sqrt{\frac{T_z}{\Omega^2}} \tag{215}$$

$$\rho(z) = \rho_0 \exp\left(-\frac{z^2}{2h^2}\right) \tag{216}$$

Thus the central density ρ_0 and the mean density $\langle\rho\rangle$ are related to the surface density in a simple way:

$$\rho_0 = \frac{\Sigma}{\sqrt{2\pi}h}$$
$$\langle\rho\rangle = \frac{\rho_0}{\sqrt{2}} \tag{217}$$

The triaxial Gaussian velocity distribution is equivalent to a Rayleigh distribution of the orbital elements e and i[21]:

$$dn(e^2, i^2) = \frac{1}{\langle e^2\rangle\langle i^2\rangle} \exp\left(-\frac{e^2}{\langle e^2\rangle} - \frac{i^2}{\langle i^2\rangle}\right) de^2 di^2$$
$$\langle e^2\rangle = \frac{2T_r}{(\Omega r_0)^2} \qquad \langle i^2\rangle = \frac{2T_z}{(\Omega r_0)^2} \tag{218}$$

Planetesimal encounters couple the time evolution of eccentricity and inclination, so that the ratio i^2/e^2 tends to an equilibrium value after a few relaxation times. It is close to 1/4 in a Keplerian potential, but the precise value also depends on the potential itself (Ida et al. 1993).

16.2 Dynamical Friction

Planetesimal–planetesimal scatterings change the velocity distribution through two different processes. Firstly, it is unlikely that two planetesimals scatter each other on circular orbits. Thus we expect a steady increase of the velocity dispersion due to this viscous stirring. Secondly, encounters between unequal masses lead successively to energy equipartition, slowing down the larger bodies through dynamical friction. The later mechanism is not related to the disc geometry at all, but operates in any multi-mass system. A special case is the systematic deceleration of a massive body M in a homogeneous sea of lighter particles m with density n_0, which is given by the Chandrasekhar dynamical friction formula (Chandrasekhar 1942)

$$\frac{d\mathbf{v_M}}{dt} = -\mathbf{v_M} \frac{4\pi \ln\Lambda G^2 (M+m) n_0 m}{v_M^3} \left(\mathrm{erf}(X) - \frac{2X}{\sqrt{\pi}} e^{-X^2}\right)$$
$$X = \frac{v_M}{\sqrt{2}\sigma_v} \tag{219}$$

where σ_v is the velocity dispersion of the lighter particles. The Coulomb logarithm Λ arises from an integration over all impact parameters smaller than an upper limit $l_{\max}$ and is given by

$$\Lambda \approx \frac{\sigma_v^2 l_{\max}}{G(m+M)} \tag{220}$$

[21] Eq. 10–12 provide the coordinate transformation.

Although encounters in the gravitational field of the sun deviate from pure two-body scatterings, it is safe to neglect the presence of the sun if the encounter velocity is large compared to the Hill velocity[22] ΩR_{Hill}. Thus the classical dynamical friction formula is also applicable to planetesimal encounters in the high velocity regime, though a generalisation to triaxial velocity distributions σ_i is necessary (Binney 1977,see e.g.):

$$\frac{dv_{M,i}}{dt} = -v_{M,i}\sqrt{2\pi}G^2 \ln(\Lambda)(M+m) n_0 m B_i \tag{221}$$

$$B_i = \int_0^\infty \exp\left(-\frac{1}{2}\sum \frac{v_j^2}{\sigma_j^2+u}\right) \frac{du}{\sqrt{(\sigma_1^2+u)(\sigma_2^2+u)(\sigma_3^2+u)}(\sigma_i^2+u)} \tag{222}$$

An additional complication is the choice of $l_{\max}$ (i. e. the choice of the Coulomb logarithm). There are several scale lengths, which could determine the largest impact parameter $l_{\max}$: The scale height of the planetesimal disc, the radial excursion due to the excentric motion of the planetesimals and the Hill radius of the planetesimals. As it is not possible to derive a unique expression for $l_{\max}$ from first principles, a proper formula is often fitted to N–body calculations (compare Eq. 235). The velocity dispersion of a planetesimal disc is triaxial with $T_\phi/T_r = 1/4$ and $T_z/T_r \approx 1/4$. We take these values and expand Eq. 221 for small velocities v_M:

$$\frac{dv_{M,r}}{dt} \approx -1.389\ v_{M,r} \frac{\sqrt{2\pi}G^2 \ln(\Lambda)(M+m) n_0 m}{T_r^{3/2}}$$
$$\frac{dv_{M,\phi}}{dt} \approx -3.306\ v_{M,\phi} \frac{\sqrt{2\pi}G^2 \ln(\Lambda)(M+m) n_0 m}{T_r^{3/2}}$$
$$\frac{dv_{M,z}}{dt} \approx -3.306\ v_{M,z} \frac{\sqrt{2\pi}G^2 \ln(\Lambda)(M+m) n_0 m}{T_r^{3/2}} \tag{223}$$

The derived expressions provide a compact tool to analyse dynamical friction in disc systems. However, the involved approximations are too severe compared to the needs of an accurate description. While these concise expressions are valuable for basic estimations, the following sections derive viscous stirring and dynamical friction formulæ for a planetesimal system in a rigorous way.

16.3 High Speed Encounters

We return to the Boltzmann equation as a starting point for the derivation of the scattering coefficients:

$$\left(\frac{\partial f}{\partial t}\right)_{\mathrm{coll}} = \frac{\partial f}{\partial t} + \mathbf{v}\cdot\nabla f - \nabla\Phi\cdot\frac{\partial f}{\partial \mathbf{v}} \tag{224}$$

In virtue of the ansatz for the distribution function (see Eq. 210), it is sufficient to derive the time derivative of the second order velocity moments T_r and T_z. Since the distribution function is time independent in the absence of encounters, only the collisional term contributes to the time derivative of the velocity dispersions T_k ($k \in (r, z, \phi)$ in the

[22] Whenever relative velocities are classified as "high" or "low" in the following sections, a comparison with the Hill velocity is implied.

following):

$$\frac{d\rho T_k}{dt} = \int d^3v m v_k^2 \left(\frac{\partial f}{\partial t}\right)_{\rm coll} \tag{225}$$

The collisional term invokes the averaging over many different scattering trajectories and is, given that the underlying encounter model is analytically solvable, still too complex to derive an exact expression. If most of the encounters are weak – a realistic assumption in a planetesimal disc – it is possible to expand the collisional contribution in terms of the velocity change Δv_i. This is the Fokker-Planck approximation (see e.g. Binney 1994))

$$\left(\frac{\partial f}{\partial t}\right)_{\rm coll} = -\sum_i \frac{\partial}{\partial v_i}[fD(\Delta v_i)] + \frac{1}{2}\sum_{i,j}\frac{\partial^2}{\partial v_i \partial v_j}[D(\Delta v_i, \Delta v_j)] \tag{226}$$

where the diffusion coefficients D contain all information on the underlying scattering process. Next we consider two interacting planetesimal populations m, m^* with distribution functions

$$f = \frac{\Omega\Sigma}{2\pi^2 T_r T_z m}\exp\left(-\frac{v_r^2+4v_\phi^2}{2T_r} - \frac{v_z^2+\Omega^2 z^2}{2T_z}\right)$$
$$f^* = \frac{\Omega\Sigma^*}{2\pi^2 T_r^* T_z^* m^*}\exp\left(-\frac{v_r^2+4v_\phi^2}{2T_r^*} - \frac{v_z^2+\Omega^2 z^2}{2T_z^*}\right) \tag{227}$$

to evaluate the terms in equation 226. We follow Stewart (2000) except some minor changes in the notation. The collisional term requires an averaging over the velocities of the two interacting planetesimals m and m^*:

$$\frac{d\langle\rho v_k^2\rangle}{dt} = 2\pi G^2 m m^{*2}\int d^3v \int d^3v^* f f^* \times \left[-\frac{2A(m+m^*)u_k v_k}{m^* u^3} + \frac{Bu^2 + (2C-B)3u_k^2}{u^3}\right]$$
$$u_k = v_k - v_k^*$$
$$A = \ln(\Lambda^2+1) \quad C = \frac{\Lambda^2}{\Lambda^2+1} \quad B = A - C \tag{228}$$

A coordinate transformation to the relative velocity u and the modified centre-of-mass velocity w

$$w_k = \begin{cases} V_k + \dfrac{(m^*T_r^* - mT_r)u_k}{(m+m^*)(T_r+T_r^*)} & \text{for } k \in \{r,\phi\} \\ V_k + \dfrac{(m^*T_z^* - mT_z)u_k}{(m+m^*)(T_z+T_z^*)} & \text{for } k = z \end{cases}$$
$$\mathbf{V} = \frac{m\mathbf{v} + m^*\mathbf{v}^*}{m+m^*} \tag{229}$$

further simplifies the double integral. Thus the integration separates in a simple integral over w and a more demanding u–integration:

$$\frac{d\langle\rho v_r^2\rangle}{dt} = 2\pi G^2 m m^{*2}\int d^3w \int d^3u f f^* \times \left[\frac{2A(m^*T_r^* - mT_r)u_r^2}{m^*(T_r+T_r^*)u^3} + \frac{B(u^2-3u_r^2)}{u^3}\right]$$
$$\frac{d\langle\rho v_\phi^2\rangle}{dt} = 2\pi G^2 m m^{*2}\int d^3w \int d^3u f f^* \times \left[\frac{2A(m^*T_r^* - mT_r)u_\phi^2}{m^*(T_r+T_r^*)u^3} + \frac{B(u^2-3u_\phi^2)}{u^3}\right]$$
$$\frac{d\langle\rho v_z^2\rangle}{dt} = 2\pi G^2 m m^{*2}\int d^3w \int d^3u f f^* \times \left[\frac{2A(m^*T_z^* - mT_z)u_z^2}{m^*(T_z+T_z^*)u^3} + \frac{B(u^2-3u_z^2)}{u^3}\right]$$

All integrals are solvable and give the result

$$\frac{d\langle v_{r,\phi}^2\rangle}{dt} = \frac{G^2\rho^*}{2\sqrt{2}(T_r+T_r^*)^{3/2}} [B(T_r^*+T_r)m^* J_{r,\phi}(\beta) + 2A(T_r^* m^* - mT_r)H_{r,\phi}(\beta)]$$
$$\frac{d\langle v_z^2\rangle}{dt} = \frac{G^2\rho^*}{2\sqrt{2}(T_r+T_r^*)^{1/2}(T_z+T_z^*)} [B(T_z^*+T_z)m^* J_z(\beta) + 2A(T_z^* m^* - mT_z)H_z(\beta)]$$
$$\beta^2 := \frac{T_z+T_z^*}{T_r+T_r^*} \tag{230}$$

where six auxiliary functions are introduced to arrive at a more compact notation:

$$a = \sqrt{4-3x^2} \quad b = \sqrt{1-(1-\beta^2)x^2}$$
$$H_r := 8\sqrt{\pi}\int_0^1 \frac{x^2}{ab}dx$$
$$H_\phi := 8\sqrt{\pi}\int_0^1 \frac{1-x^2}{a(\beta a+b)}dx$$
$$H_z := 8\sqrt{\pi}\int_0^1 \frac{\beta(1-x^2)}{b(\beta a+b)}dx$$
$$J_r := -2H_r + H_\phi + H_z$$
$$J_\phi := \quad H_r - 2H_\phi + H_z$$
$$J_z := \quad H_r + H_\phi - 2H_z \tag{231}$$

Since these are non-trivial functions, we apply a standard Chebyshev approximation for $\beta \in [0,1]$:

$$f(x) \approx \sum_{k=0}^{5} c_k T_k(x) - \frac{1}{2}c_0 \tag{232}$$

Table 9 summarises the Chebyshev coefficients. A final z–averaging yields the expressions:

$$\frac{d\langle v_{r,\phi}^2\rangle}{dt} = \frac{G^2\Omega\Sigma^*}{4\sqrt{\pi}(T_r+T_r^*)^{3/2}(T_z+T_z^*)^{1/2}} \times [B(T_r^*+T_r)m^* J_{r,\phi}(\beta) + 2A(T_r^* m^* - mT_r)H_{r,\phi}(\beta)] \tag{233}$$
$$\frac{d\langle v_z^2\rangle}{dt} = \frac{G^2\Omega\Sigma^*}{4\sqrt{\pi}(T_r+T_r^*)^{1/2}(T_z+T_z^*)^{3/2}} \times [B(T_z^*+T_z)m^* J_z(\beta) + 2A(T_z^* m^* - mT_z)H_z(\beta)] \tag{234}$$

The determination of a proper Coulomb logarithm Λ leaves room for further optimisation. A careful comparison with N–body models gives rise to the empirical choice (Ohtsuki

f	J_r	J_ϕ	J_z	H_r	H_ϕ	H_z
c_0	-10.34660733	1.81674741	8.52985992	11.00434580	6.94989422	4.71219005
c_1	4.69990443	2.95397208	-7.65387651	-2.64707927	-2.06510182	1.47084771
c_2	-1.25533220	-1.18724874	2.44258094	0.60969641	0.58700192	-0.62294130
c_3	0.30288875	0.37775788	-0.68064662	-0.13815856	-0.16311494	0.18968657
c_4	-0.07040537	-0.11070339	0.18110876	0.03112047	0.04455314	-0.05271757
c_5	0.01540098	0.02922947	-0.04463045	-0.00669979	-0.01130929	0.01331068
Δ	0.006	0.015	0.022	0.0025	0.0058	0.0066

Table 9. Chebyshev coefficients of the auxiliary functions J_k and H_k.

et al. 2002):

$$\Lambda = \frac{1}{12}(\langle\tilde{e}^2\rangle + \langle\tilde{i}^2\rangle)\langle\tilde{i}^2\rangle^{1/2} \tag{235}$$

$$\tilde{e} = \frac{\sqrt{2T_r}}{\Omega R_{\rm Hill}} \qquad \tilde{i} = \frac{\sqrt{2T_z}}{\Omega R_{\rm Hill}} \tag{236}$$

Ohtsuki et al. 2002 also report a further improvement by setting $B \equiv A$.

16.4 Low Speed Encounters

Encounters in the low velocity regime exhibit a wealth of different orbits, as the solar gravity field perturbs the two-body scattering. Only a small subset of the trajectories represents simple, regular orbits like Tadpole or Horseshoe orbits[23]. Hence an examination of this velocity regime is done best with a numerical study of the parameter space by integrating the equations of motions numerically (see Eq. 16).

Ohtsuki et al. (2002) integrated a large set of planetesimal encounters and extracted fitting formulæ that cover the low velocity regime. Their expressions for viscous stirring are:

$$\begin{aligned}
\frac{dT_r}{dt} &= \frac{Gr_0\Omega h\Sigma^*}{6(m+m^*)} 73 C_1 m^* \\
\frac{dT_z}{dt} &= \frac{Gr_0\Omega h\Sigma^*}{6(m+m^*)} C_2 m^* \left(4\langle\tilde{i}^2\rangle + 0.2(\langle\tilde{e}^2\rangle)^{3/2}\sqrt{\langle\tilde{i}^2\rangle}\right) \\
\tilde{e} &= e/h \qquad \tilde{i} = i/h \\
C_1 &:= \frac{\ln(10\Lambda^2/\tilde{e}^2+1)}{10\Lambda^2/\tilde{e}^2} \\
C_2 &:= \frac{\ln(10\Lambda^2\sqrt{\tilde{e}^2}+1)}{10\Lambda^2\sqrt{\tilde{e}^2}}
\end{aligned} \tag{237}$$

The stirring rate of the radial velocity dispersion approaches a finite value for very low velocity dispersions, while the stirring rate for the vertical velocity dispersion drops to zero as the velocity dispersion decreases. This different behaviour of the two limits is due to the encounter geometry: If two planetesimals have zero inclination, they may still excite higher eccentricities during an encounter, but they remain confined to the initial orbital plane preventing any excitation of inclinations.

The respective expressions for the dynamical friction rates are:

$$\begin{aligned}
\frac{dT_r}{dt} &= \frac{Gr_0\Omega h\Sigma^*}{6(m+m^*)(T_r+T_r^*)} 10 C_3 \langle\tilde{e}^2\rangle (T_r^* m^* - T_r m) \\
\frac{dT_z}{dt} &= \frac{Gr_0\Omega h\Sigma^*}{6(m+m^*)(T_z+T_z^*)} 10 C_3 \langle\tilde{i}^2\rangle (T_z^* m^* - T_z m) \\
C_3 &:= \frac{\ln(10\Lambda^2+1)}{10\Lambda^2}
\end{aligned} \tag{238}$$

As the stirring rates are only valid in the low velocity regime, Ohtsuki et al. (2002) introduced special interpolation coefficients C_i. These coefficients tend to unity for very small velocity dispersions, and drop to zero in the high velocity regime. Thus the interpolation formulæ are properly "switched off" in the high velocity regime, so they do not interfere with the known high velocity stirring rates.

16.5 Distant Encounters

All formulæ include only the stirring rates due to close encounters, but non-crossing orbits also contribute to the overall change of the velocity distribution. As these distant encounters lead to small changes of the orbital elements, the problem is accessible to perturbation theory; see Hasegawa (1990) for a detailed treatment. Stewart (2000) integrated the perturbation solution over all impact parameters to derive the collective effect of all distant encounters:

$$\begin{aligned}
\frac{d\langle e^2\rangle}{dt} &= \frac{\Omega m^* \Sigma^* r_0^2}{(m+m^*)^2} \langle P_{\rm VS,dist}\rangle \\
\langle P_{\rm VS,dist}\rangle &= 7.6 \frac{\alpha(m+m^*)^2}{M_c^2} \\
&\quad \frac{\mathrm{EXINT}\left(\alpha\frac{h^2}{(\langle e^2\rangle+\langle e^{*2}\rangle)}\right) - \mathrm{EXINT}\left(\alpha\frac{h^2}{(\langle i^2\rangle+\langle i^{*2}\rangle)}\right)}{\langle e^2\rangle + \langle e^{*2}\rangle - \langle i^2\rangle - \langle i^{*2}\rangle} \\
\mathrm{EXINT}(x) &:= \exp(x)\Gamma(0,x) \qquad h = \sqrt[3]{\frac{m+m^*}{3M_c}} \\
&\alpha \approx 1
\end{aligned} \tag{239}$$

α accounts for the uncertainty in the smallest impact parameter that is regarded as a distant encounter. While distant encounters are already included in the interpolation formula

[23] The most famous example of such a regular orbit are the two saturnian moons Janus and Epimetheus which share nearly the same orbit.

of the low–velocity regime, we use the modified expression:

$$\left(\frac{dT_r}{dt}\right)_{\rm dist} = \frac{1}{2}(\Omega r_0)^2 \left(\frac{d\langle e^2\rangle}{dt}\right)_{\rm dist}(1-C_1) = \frac{GM_c r_0 \Omega m^* \Sigma^*}{2(m+m^*)^2}\langle P_{\rm VS,dist}\rangle(1-C_1) \quad (240)$$

Stewart (2000) omitted the change in the inclination, as it is small due to the encounter geometry. Nevertheless we derived the integrated stirring rate for completeness, which we give in 241 A close inspection of the integrated perturbation shows that the above formula is roughly a factor $\langle i^2\rangle + \langle i^{*2}\rangle$ smaller than the corresponding changes in the eccentricity.

16.6 Gas Damping

The presence of a gaseous disc damps the velocity dispersion of the planetesimals and introduces a slow inward migration. Adachi et al. (1976) used the drag law Eq. 75 to approximate[24] the average change of the orbital elements:

$$\tau_0 = \frac{2m}{\pi C_D \rho_g R^2 v_K} \qquad \eta_g = \frac{|v_K - v_g|}{v_K}$$

$$\frac{d}{dt}e^2 \approx -\frac{2e^2}{\tau_0}\left(0.77\,e + 0.64\,i + \frac{3}{2}\eta_g\right) \quad (242)$$

$$\frac{d}{dt}i^2 \approx -\frac{2i^2}{\tau_0}\left(0.39\,e + 0.43\,i + \frac{1}{2}\eta_g\right)$$

$$\frac{d}{dt}a \approx -\frac{2a}{\tau_0}\eta_g\,(0.97\,e + 0.64\,i + \eta_g) \quad (243)$$

η_g is the dimensionless velocity lag of the sub–Keplerian rotating gaseous disc.

16.7 Unified Expressions

All expressions for the different velocity regimes are constructed such that a smooth transition between the different regimes is assured. Thus, a simple addition of all contributions yields already the unified expressions

$$\frac{dT_r}{dt} = \left(\frac{dT_r}{dt}\right)_{\rm high} + \left(\frac{dT_r}{dt}\right)_{\rm low} + \left(\frac{dT_r}{dt}\right)_{\rm gas} + \left(\frac{dT_r}{dt}\right)_{\rm dist} \quad (244)$$

$$\frac{dT_z}{dt} = \left(\frac{dT_z}{dt}\right)_{\rm high} + \left(\frac{dT_z}{dt}\right)_{\rm low} + \left(\frac{dT_z}{dt}\right)_{\rm gas} + \left(\frac{dT_z}{dt}\right)_{\rm dist} \quad (245)$$

which cover the full range of relative velocities. Although only two populations m and m^* were assumed, Eq. 244 and Eq. 245 are readily generalised to a multi-mass system by adding a summation over all masses.

16.8 Inhomogeneous Disc

The preceding derivations assumed a homogeneous disc, which simplified the calculation, since the integration over all impact parameters needed no special precaution. A more sophisticated consequence is that the spatial density and the density in semimajor axis space are equal:

$$\Sigma(r) = \Sigma(a) = \Sigma_0 \quad (246)$$

Density inhomogeneities break this simple relation, as particles at the same radial distance could have different semimajor axes, and particles with the same semimajor axis are located at different positions at a given time. While both representations are equivalent (i. e. describe the same system in different ways), we chose the density in semimajor axis space as the primary density[25]. The spatial density is derived as:

$$\Sigma(r) = \int \frac{1}{\sqrt{2\pi a^2\langle e^2(a)\rangle}} \exp\left(-\frac{(a-r)^2}{2a^2\langle e^2(a)\rangle}\right)\Sigma(a)da \quad (247)$$

Likewise, T_r and T_z are also functions of the semimajor axis.

Furthermore, an inhomogeneous surface density invalidates the averaging over all impact parameters. Planetesimal encounter are most efficient for impact parameters smaller than a few Hill radii, so the derivation is still valid if the surface density is roughly constant on that length scale. However, a planetesimal that is large enough will "feel" the spatial inhomogeneities or even generates density fluctuations. Hence it is essential to extend the validity of the averaged expressions to inhomogeneous systems. We use the averaged expressions

$$\left\langle \frac{dT_{r,z}}{dt}\right\rangle = \Sigma(a)\int_{-\infty}^{\infty}\frac{d\tilde{T}_{r,z}(b)}{dt}db \quad (248)$$

as a starting point ($d\tilde{T}_{r,z}/dt$ excludes the surface density, as opposed to the averaged expressions). The (yet unknown) scattering contribution $d\tilde{T}_{r,z}/dt$ as a function of the impact parameter b is our starting point for a general expression for a varying surface density:

$$\frac{dT(a_0)_{r,z}}{dt} = \int_{-\infty}^{\infty}\Sigma(a_0+b)\frac{d\tilde{T}_{r,z}(b)}{dt}db \quad (249)$$

We restate Eq. 249 in terms of a weight function $w(b)$:

$$\frac{dT(a_0)_{r,z}}{dt} = \left\langle\frac{dT_{r,z}}{dt}\right\rangle\frac{1}{\Sigma(a_0)}\int_{-\infty}^{\infty}\Sigma(a_0+b)w(b)db \quad (250)$$

$$w(b) = \Sigma(a_0)\frac{d\tilde{T}_{r,z}(b)}{dt}\left\langle\frac{dT_{r,z}}{dt}\right\rangle^{-1} \quad (251)$$

The numerical solution of the Hill problem (see Eq. 16) gives some insight into how the weight function $w(b)$ changes with the impact parameter. Fig. 12 illustrates the change in e^2 of the relative motion during an encounter of two planetesimals that were initially on circular orbits. While the details depend on the initial inclination and eccentricity as well as on the selected orbital element, all result share some basic features. Small (compared to the Hill radius) impact parameters allow for a horseshoe orbit and the change in the orbital elements is small except a change in the semimajor axis. Intermediate impact parameters which lead to close encounters provide the strongest perturbation, but they are also more susceptible to complicated dynamics (compare the

[24] A formal expansion at $e = 0$, $i = 0$, $\eta_g = 0$ is not possible, since the drag law involves the modulus of the relative velocity. Kary et al. (1993) corrected a missing factor 3/2 in Eq. 242.

[25] We denote $\Sigma(a)$ also as "surface density" and refer to a as a radial coordinate. However, all formulæ are precise in discriminating both representations in r and a.

$$\frac{d\langle i^2\rangle}{dt} = \frac{\Omega m^* \Sigma^* r_0^2}{(m+m^*)^2}\langle Q_{\rm VS,dist}\rangle\langle Q_{\rm VS,dist}\rangle = 0.4\frac{\alpha^2(m+m^*)^2}{M_c^2} \times \frac{1}{\langle e^2\rangle + \langle e^{*2}\rangle - \langle i^2\rangle - \langle i^{*2}\rangle} \times \left[1 - \frac{\alpha h^2}{\langle i^2\rangle + \langle i^{*2}\rangle}\mathrm{EXINT}\left(\alpha\frac{h^2}{(\langle i^2\rangle + \langle i^{*2}\rangle)}\right) - (\langle i^2\rangle + \langle i^{*2}\rangle)\frac{\mathrm{EXINT}\left(\alpha\frac{h^2}{(\langle e^2\rangle + \langle e^{*2}\rangle)}\right) - \mathrm{EXINT}\left(\alpha\frac{h^2}{(\langle i^2\rangle + \langle i^{*2}\rangle)}\right)}{\langle e^2\rangle + \langle e^{*2}\rangle - \langle i^2\rangle - \langle i^{*2}\rangle}\right] \quad (241)$$

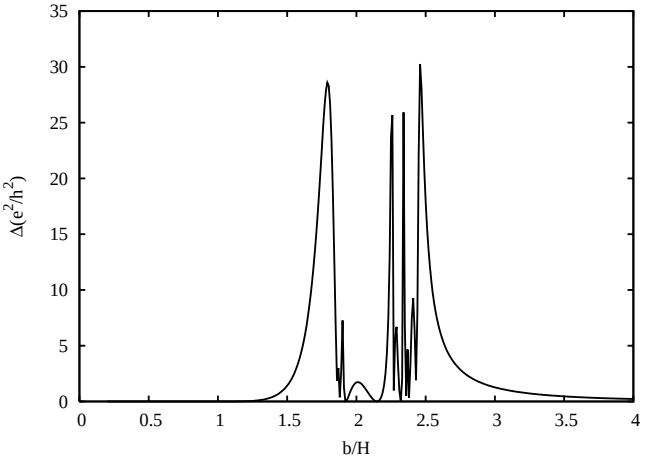

Figure 12. Change of the relative eccentricity e^2 due to an encounter of two bodies initially on circular orbits. b/H is the impact parameter in units of the Hill radius. The plot was obtained by integrating Eq. 16.

resonant structures in Fig. 12). As the gravitational attraction drops with increasing distance, non-crossing orbits yield ever smaller perturbations with increasing impact parameter. Aside from this qualitative behaviour, it is very difficult to derive precise expressions. While the limit of high velocities is accessible through the two-body approximation, any general formula involves some empiric interpolation to cover the full parameter space (Rafikov 2003b,a,see the approximations of). Therefore we decided to approximate the weight function such that the main features of the true weight function $w(b)$ are reproduced. While this approach is less accurate, it provides better insight into the involved approximations. We expand the surface density under the integral in Eq. 250 and compare the expansion coefficients for $w(b)$ and the approximation $\tilde{w}(b)$ to derive constraints on the choice of $\tilde{w}(b)$. The lowest non-vanishing order is:

$$l^2 = \int_{-\infty}^{\infty} b^2 w(b) db = \int_{-\infty}^{\infty} b^2 \tilde{w}(b) db \quad (252)$$

l can be interpreted as the width of the heating zone. Condition 252 inspires our choice of the weight function $\tilde{w}(b)$

$$\tilde{w}(b) = \frac{1}{\sqrt{2\pi}l}\exp\left(-\frac{b^2}{2l^2}\right)$$
$$l^2 = \frac{1}{\Omega^2}(T_r^{(i)} + T_r^{(j)}) + R_{\rm Hill}^2 \quad (253)$$

where $T_r^{(i)}$ and $T_r^{(j)}$ are the radial velocity dispersions of the interacting radial bins and l is adjusted to the findings of Ida (1993). The advantage of the bell curve is that it has a discrete counterpart

$$\tilde{w}(b) \approx \frac{1}{\Delta a 2^N}\binom{N}{b/\Delta a + N/2} \qquad N = 4l^2/(\Delta a)^2 \quad (254)$$

which makes the weight function readily applicable to the summation on an equidistant radial grid with spacing Δa.

16.9 Diffusion Coefficient

We concentrated on the evolution of the velocity dispersion so far, but scatterings among planetesimals also change the semimajor axis of the disc particles, inducing a diffusive evolution of the surface density:

$$\frac{\partial\Sigma}{\partial t} = \Delta_a(D\Sigma) \quad (255)$$

The diffusion coefficient D is related to the typical change in semimajor axis Δa and the timescale $T_{\rm 2Body}$ on which planetesimal encounters operate:

$$D \approx \frac{(\Delta a)^2}{T_{\rm 2Body}} \quad (256)$$

If we neglect the radial displacement during an encounter, the change in semimajor axis is solely due to the change of the velocity:

$$-\frac{GM}{2a} = -\frac{GM}{r} + \frac{1}{2}v^2$$
$$\Delta a \approx \frac{2a^2}{GM}\mathbf{v}\cdot\Delta\mathbf{v} \quad (257)$$

An average over all orientations of the velocity $\mathbf{v}$ and the velocity change $\Delta\mathbf{v}$ yields the mean square change in semimajor axis:

$$\langle(\Delta a)^2\rangle \approx \frac{4a^3}{3GM}\langle(\Delta v)^2\rangle$$
$$= \frac{4}{3\Omega^2}(\Delta T_r + \Delta T_\phi + \Delta T_z) \quad (258)$$

This yields the mean diffusion coefficient

$$D \approx \frac{4}{3\Omega^2}\left(\frac{5}{4}\frac{d}{dt}T_r + \frac{d}{dt}T_z\right) \quad (259)$$

where the time derivatives of the velocity dispersions T_r and T_z are taken with respect to encounters.

16.10 Coagulation Equation

We already stated the coagulation equation for a multi-mass system:

$$0 = \frac{\partial}{\partial t}mn(t,m) + \frac{\partial}{\partial m}F_m(t,m) \quad (260)$$
$$F_m = -\iint n_1(t,m_1)n_2(t,m_2)\sigma(m_1,m_2)v_{\rm rel}$$
$$M_{\rm red}(m,m_1,m_2)dm_1dm_2 \quad (261)$$

Since the vertical density profile of a planetesimal disc is specified by the known distribution function, we insert the

isothermal density profile (see Eq. 216) in the coagulation equation 260 and integrate over z:

$$0 = \frac{\partial}{\partial t}\Sigma(t,m) + \frac{\partial}{\partial m}F_m(t,m) \tag{262}$$

$$F_m = -\iint \frac{1}{\sqrt{2\pi(h(m_1)^2+h(m_2)^2)}}\frac{\Sigma_1(t,m_1)}{m_1}\frac{\Sigma_2(t,m_2)}{m_2} \tilde{\sigma}(m_1,m_2)v_{\rm rel}M_{\rm red}(m,m_1,m_2)dm_1dm_2 \tag{263}$$

$\Sigma(m)$ is a short-hand notation for the differential surface density $\frac{d\Sigma}{dm}$. Further integration over all masses gives the total mass balance:

$$\Sigma_{\rm tot} = \int_{m_{\rm min}}^{m_{\rm max}} \Sigma(t,m)dm$$

$$\frac{\partial}{\partial t}\Sigma_{\rm tot} = F_m(m_{\rm min}) - F_m(m_{\rm max}) \tag{264}$$

The calculation of collisional cross sections is closely related to the underlying encounter dynamics. A homogenous system introduces no systematic perturbation, hence an encounter is a pure two-body problem which is analytically solvable. Thus it is possible to derive the cross section without any approximation. Since encounters in the field of a central star deviate noticeably from the pure Kepler solution, the cross sections are also modified. While the cross section in the high velocity regime reduces to the two-body formula (except minor corrections), the low velocity regime is explored best by numerical calculations. It is not appropriate to disentangle the different contributions in Eq. 263, but to combine the various terms to the collisional probability

$$P_{\rm coll} = \frac{\tilde{\sigma}(m_1,m_2)v_{\rm rel}}{\sqrt{2\pi(h(m_1)^2+h(m_2)^2)}} \tag{265}$$

which can be easily deduced from the fraction of colliding orbits in Monte–Carlo simulations. An accurate expression for the collisional probability should include the two-body cross section in the limit of high velocities and the numerical data for the low velocity regime as well. We use numerical calculations from Greenberg (1991); Greenzweig (1992) [26] as a basis for a unified fitting formula

$$\tilde{\sigma} = \sigma \times 0.572\,(1+3.67v_{\rm Hill}/v_{\rm rel})\left(1+1.0\frac{\sigma\Omega^2}{T_z}\right)^{-1/2} \tag{266}$$

$$\sigma = \sigma_{\rm geom}\left(1+\frac{v_\infty^2}{v_{\rm rel}^2+1.8v_{\rm Hill}^2}\right) \tag{267}$$

$$v_{\rm rel}^2 = \frac{1}{2}(T_r+T_\phi+T_z) \qquad v_{\rm Hill} = \Omega r_{\rm Hill} \tag{268}$$

which gives an effective cross section $\tilde{\sigma}$ for planetesimal–planetesimal encounters. Eq. 266 reduces to the well-known gravitational focusing formula in the limit of high velocities:

$$\tilde{\sigma} \propto \sigma_{\rm geom}\left(1+\frac{v_\infty^2}{v_{\rm rel}^2}\right) \tag{269}$$

$$v_\infty^2 = \frac{2G(m_1+m_2)}{R_1+R_2} \tag{270}$$

If the vertical velocity dispersion is small, the disc becomes two-dimensional and the cross section is proportional to R. The main differences to the two-body cross section 269 is a finite gravitational focusing factor, since the Keplerian shear inhibits a zero relative velocity, and a finite collisional probability for very small velocities, again due to the shear which provides a finite influx of particles.

The precise calculation of the coagulation kernel should include an integration over all semimajor axes with a proper weighting kernel. As collisions among particles in the statistical model play only a major role when the system is still homogenous, we omitted this contribution. In addition, this helps saving computational time, since the solution of the coagulation equation is very costly. However, interactions between N–body particles and the statistical model include spatial inhomogeneities properly (see section 17).

16.11 Collisional Damping

Collisions are a dissipative process that removes kinetic energy from the planetesimal system and damps the eccentricity. Low speed encounters leave the colliding bodies intact and damp the relative velocities through inelastic collisions. In contrast, high velocity encounters disrupt the colliding bodies and turn them into an expanding cloud of fragments. As a major part of the initial kinetic energy is converted into heat, the fragments disperse with rather low velocities thus reducing the overall velocity dispersion. We formulate the dissipation due to collisions analogue to Eq. 262:

$$0 = \frac{\partial}{\partial t}T_k\Sigma(t,m) + \frac{\partial}{\partial m}F_{Q,k}(t,m)$$

$$F_{Q,k} = -\iint \frac{1}{\sqrt{2\pi(h(m_1)^2+h(m_2)^2)}}\frac{\Sigma_1(t,m_1)}{m_1}\frac{\Sigma_2(t,m_2)}{m_2} \tilde{\sigma}(m_1,m_2)v_{\rm rel}Q_{{\rm red},k}(m,m_1,m_2)dm_1dm_2 \qquad k\in\{r,z\} \tag{271}$$

$Q_{{\rm red},k}$ is the kinetic energy redistribution function and $F_{Q,k}$ is the associated flux across the mass distribution. k indicates the two velocity dispersions. $Q_{{\rm red},k}$ is a complex function, since the disruption of a planetesimal produces fragments with a large scatter in velocities and a complicated velocity field. The velocity of a fragment consists of two contributions: The ejection velocity relative to the target and the velocity $\bar{v}$ of the target within the corotating coordinate system. Owing to the strong dissipation, fragment velocities are dominated by the motion of the centre-of-mass of the two colliding bodies. Thus we neglect the ejection velocities and estimate the centre-of-mass motion. The initial kinetic energy of two colliding bodies is:

$$E_{\rm ini} = \frac{1}{2}m_1v_1^2 + \frac{1}{2}m_2v_2^2 \tag{272}$$

We separate $E_{\rm ini}$ into centre-of-mass motion and relative motion and average over $\mathbf{v}_1, \mathbf{v}_2$:

$$E_{\rm ini} = \frac{1}{2}M\left(\frac{m_1\mathbf{v}_1+m_2\mathbf{v}_2}{M}\right)^2 + \frac{1}{2}\mu(v_1^2+v_2^2)$$

$$\langle E_{\rm ini}\rangle = \frac{1}{2M}\left(m_1^2T_1+m_2^2T_2\right) + \frac{1}{2}\mu(T_1+T_2) \tag{273}$$

Most of the relative kinetic energy is dissipated during the collision, so we neglect the relative motion after the incident. The final energy is

$$\frac{\langle E_{\rm final}\rangle}{\langle E_{\rm ini}\rangle} = \frac{m_1^2T_1+m_2^2T_2}{M(m_1T_1+m_2T_2)} \tag{274}$$

[26] Their work includes an averaging over the Rayleigh distributed inclinations and eccentricities of the colliding planetesimals.

which gives the drift motion $\bar{v}$ of the expanding fragment cloud:

$$\bar{v}^2 = \frac{2}{M}\langle E_{\rm final}\rangle \tag{275}$$

$Q_{\rm red}$ is therefore coupled to the fragment redistribution function $M_{\rm red}$

$$Q_{\rm red} = \bar{v}^2 M_{\rm red} + Q_{\rm diss} \tag{276}$$

where the additional function $Q_{\rm diss}$ removes the dissipated energy.

16.12 Correlation

The statistical model of a planetesimal disc does not only require a large particle number to assure a proper description of the system by a distribution function, but also the uncorrelated motion of the planetesimals. Each of the formulæ derived before involves the averaging over different impact parameters to some extend, in combination with the vital assumption that all distances are equally probable. As long as all particles are subjected to perturbations by surrounding bodies, strong correlations are suppressed. This applies to the early stages, but the formation of protoplanets introduces a few dominant bodies that are not susceptible to the perturbation of the field planetesimals. Orbit repulsion gives rise to a regular spacing of the protoplanets, which prevents mutual collisions. Therefore not all impact parameters are equally probable due to this strong correlation. Hence a statistical model is inherently not applicable to the late stages of protoplanetary growth.

Since statistical models are superior to N–body calculations with respect to speed and (effective) particle number, modifications have been proposed to remedy this problem.

The statistical model by Wetherill (1993) uses the following solution: A gravitational range Δa (or buffer zone) is attached to each planetesimal, which represents the minimal spacing that allows for stable orbits. They propose the expression

$$\Delta a = f_\Delta R_{\rm Hill} + \sqrt{2T_r/\Omega^2} \tag{277}$$

where f_Δ is the minimal spacing in terms of the Hill radius. The value of f_Δ is adopted from Birn (1973), who derived the minimal spacing that allows for stable orbits:

$$f_\Delta = 2\sqrt{3} \tag{278}$$

Thus it is possible to define a minimum mass $m_{\rm sep}$ by the assumption that all bodies larger than this critical mass maintain a clear buffer zone:

$$f = \int_{m_{\rm sep}}^{\infty} \frac{d\Sigma}{dm}\frac{2\pi a \Delta a}{m} dm$$
$$f = 1 \tag{279}$$

f is the area covered by the buffer zones (overlapping is not taken into account, therefore $f > 1$ is possible), normalised to the ring area. Planetesimals smaller than $m_{\rm sep}$ can not enforce a minimum distance to their neighbours, as the whole disc surface is already covered by the buffer zones of the largest bodies. Owing to the regular spacing introduced by the buffer zones, planetesimals larger than the critical mass are not allowed to collide with each other. This approach has also been employed by Inaba et al. (2001), who adopted

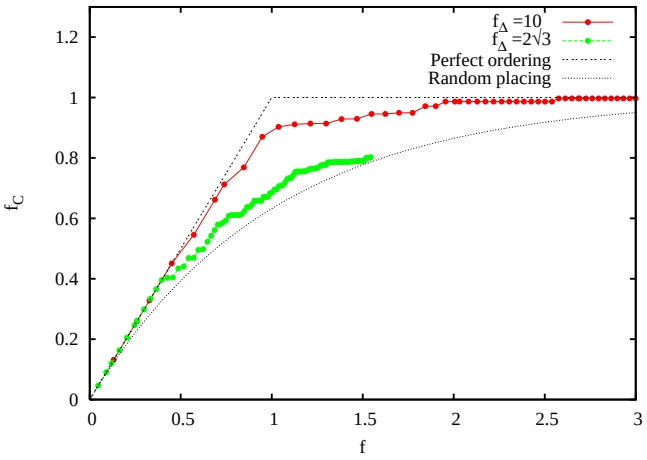

Figure 13. Covered fraction f_C as a function of f for simulation S1FB at $T = 10^5$ years. The protoplanets are already formed and grow oligarchically.

$f_\Delta = 10$, which is the mean distance of protoplanets according to the orbit repulsion mechanism.

To shed light on the proper gravitational range and the validity of this approach, we defined an additional quantity f_C which is the true area (i. e. overlapping is handled properly) covered by the buffer zones in terms of the total area:

$$f_C = \int_{m_{\rm sep}}^{\infty} \frac{d\Sigma}{dm}\frac{2\pi a(\Delta a)_C}{m} dm$$
$$f_C \leqslant 1 \tag{280}$$

Fig. 13 shows the covered fraction $f_C(m)$ as a function of the integrated buffer zones $f(m)$ for one of our hybrid calculations. Both values for f_Δ are included as well as the two limiting cases random placing and perfect ordering. Though we tested also other values of f_Δ, a spacing of ten Hill radii proved to be the best choice.

Our own experience with this method indicates that it works reasonably well and agrees with Inaba et al. (2001) who used the same technique. However, this modification includes the regular spacing of the protoplanets in a prescribed way, so any exploration of later stages, like the initiation of orbit crossing, is not accessible through this approach. Therefore we use it for comparison purposes only, since the hybrid code (see section 17) provides a much more general framework.

16.13 Discretisation

All involved quantities are only functions of a and m. Therefore we introduce a two dimensional grid, where Σ, T_r and T_z are cell centred quantities. Fig. 14 summarises the definition of the two-dimensional grid. Since the full planetesimal size range covers several orders of magnitude in mass, we chose a logarithmically equidistant discretisation in mass to cover the necessary mass range in a reliable way. The radial spacing of the grid cells is equidistant. Thus the grid setup for the mass discretisation reads (N grid cells from

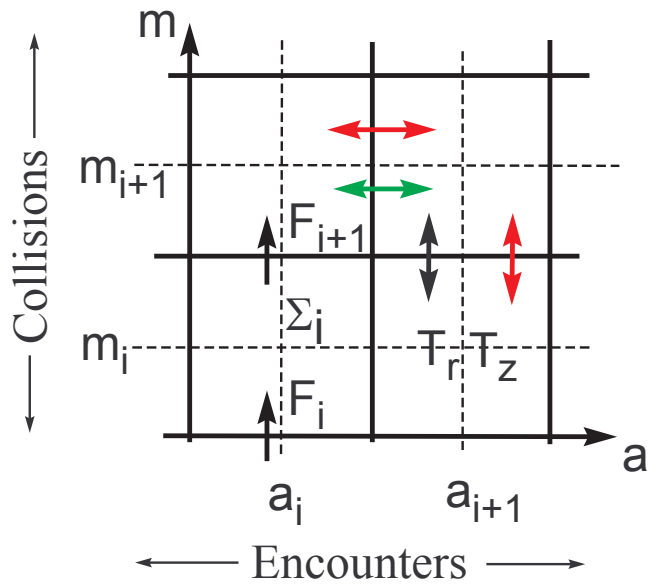

Figure 14. Numerical grid. The arrows indicate transport of kinetic energy (red), spatial transport of mass(green) and accretion (black). Non-neighbouring cells are coupled by the coagulation kernel and the radial interpolation kernel.

$m_{\min} \ldots m_{\max}$):

$$\begin{aligned} m_i &= m_{\min}\delta^{-i}(1/2+\delta/2) \qquad i = 1,\ldots,N \\ \Delta m_i &= m_{\min}\delta^{-i}(1-\delta) \\ \Sigma_i &= \frac{d\Sigma}{dm}\Delta m_i \\ \delta &= \left(\frac{m_{\min}}{m_{\max}}\right)^{1/N} \end{aligned} \tag{281}$$

The grid spacing δ controls the number of cells which are necessary to cover a specified mass range. As the evaluation of the coagulation equation scales with the third power of the number of grid cells, δ should be as large as possible. If the flux integral (see Eq. 260) is approximated in a standard way

$$\begin{aligned} F_i &= -\sum_{j=1}^{N}\sum_{k=1}^{N} F_i^{(jk)} \\ F_i^{(jk)} &= \frac{1}{\sqrt{2\pi(h_j^2+h_k^2)}}\frac{\Sigma_j}{m_j}\frac{\Sigma_k}{m_k} \times \sigma(m_j,m_k) \\ &\quad v_{\rm rel} M_{\rm red}(m_i - \Delta m_i/2, m_j, m_k) \end{aligned} \tag{282}$$

a spacing δ much smaller than 2 is required to guarantee a sufficient accuracy[27]. However, it is possible to use a spacing of 2 if special precaution is taken. Spaute et al. (1991) approximated the surface density with a power law, thus taking the gradient with respect to mass into account. While they reached only a sufficient accuracy with further special adaptations, we use a more rigorous approach. A large spacing δ reduces the accuracy, since the partial flux (Eq. 282) is strongly varying even inside one grid cell. Thus we rearrange this expression to identify the most important terms, as we can see in Eq. 283.

The strongest varying contribution F_V is now approximated by a power law with respect to m_j:

$$F_V(m) \approx F_V(m_j)\left(\frac{m}{m_j}\right)^q \tag{284}$$

Thus Eq. 284 is used to provide improved partial fluxes $F_i^{(jk)}$ and thus we come to Eq. 285.

Since the fragment redistribution function is a piecewise power law, an analytical solution of the integral is possible. Eq. 285 gives reliable results even with a spacing $\delta = 2$. The time derivative of the surface density reads

$$\dot{\Sigma}_m^i = -F_{i+1} + F_i \tag{286}$$

which assures the conservation of mass within numerical accuracy.

16.14 Integrator

All contributions to the evolution of the surface density Σ and the velocity dispersions T_r and T_z are summarised by the following set of differential equations:

$$\begin{aligned} D &= \frac{4}{3\Omega^2}\left(\frac{5}{4}\dot{T}_{r,\rm enc} + \dot{T}_{z,\rm enc}\right) \\ \frac{d\Sigma}{dt} &= \Delta_a(D\Sigma) + \dot{\Sigma}_{\rm coll} \\ \frac{d\Sigma T_r}{dt} &= \Delta_a(D\Sigma T_r) + \Sigma\dot{T}_r + \frac{d}{dt}(\Sigma T_r)_{\rm coll} \\ \frac{d\Sigma T_z}{dt} &= \Delta_a(D\Sigma T_z) + \Sigma\dot{T}_z + \frac{d}{dt}(\Sigma T_z)_{\rm coll} \end{aligned} \tag{287}$$

The Laplace operator is approximated in accordance with the equidistant radial grid:

$$\begin{aligned} \Delta_a f &= \frac{1}{a}\frac{\partial}{\partial a}\left(a\frac{\partial}{\partial a}f\right) \\ \Delta_a f &\approx \frac{f_{i+1}(1+\Delta a/(2a_i)) - 2f_i + (1-\Delta a/(2a_i))f_{i-1}}{(\Delta a)^2} \end{aligned} \tag{288}$$

We chose the Heun method[28] as the basic integrator for the statistical model. It is a second order accurate predictor–corrector scheme (X is a vector containing all the above quantities):

$$\begin{aligned} \frac{dX}{dt} &= f(X) \\ X^p &= X_n + \Delta t f(X_n) \\ X_{n+1} &= X^p + \frac{1}{2}\Delta t(f(X^p) - f(X_n)) + \mathcal{O}(\Delta t^3) \end{aligned} \tag{289}$$

The Heun method is readily extended to an iterate scheme, which is equivalent to the implicit expression:

$$X_{n+1} = X_n + \frac{1}{2}\Delta t(f(X_{n+1}) + f(X_n)) \tag{290}$$

This adds stability to the method and allows the secure integration of stiff configurations that may appear during the

[27] Ohtsuki et al. (1990) give a thorough analysis of the importance of the resolution.

[28] The name of this method is not unique. Some texts denote it as the modified Euler method. The Heun method is a special case of the Runge–Kutta methods.

$$F_i^{(jk)} = \frac{v_{\rm rel}}{\sqrt{2\pi(h_j^2+h_k^2)}} \frac{\Sigma_k}{m_k} \underbrace{\frac{\Sigma_j}{m_j} \times \sigma(m_j,m_k) m_j}_{F_V(m_j)} \frac{1}{m_j} M_{\rm red}(m_i - \Delta m_i/2, m_j, m_k) \tag{283}$$

$$j \geqslant k$$

$$F_i^{(jk)} = \frac{v_{\rm rel}}{\sqrt{2\pi(h_j^2+h_k^2)}} \frac{\Sigma_k}{m_k} \int_{m_j-\Delta m_j/2}^{m_j+\Delta m_j/2} \frac{F_V(m)}{m \Delta m_j} M_{\rm red}(m_i - \Delta m_i/2, m, m_k) dm \tag{285}$$

$$j \geqslant k$$

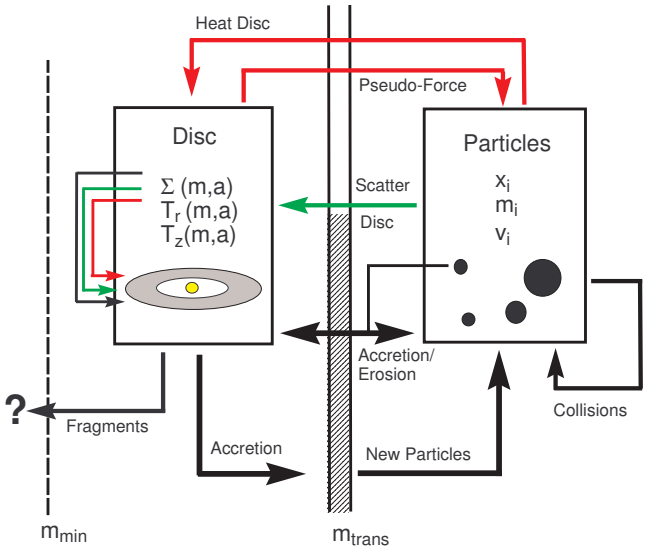

Figure 15. Interplay between the N–body component and the statistical component of the hybrid code. Black arrows indicate mass transfer, red arrows exchange of kinetic energy and green arrows indicate spatial structuring, respectively.

runaway accretion phase. In practice, three iterations are sufficient to guarantee a stable integration. As the diffusive part is discretised with a first order accurate formula (see Eq. 288), the whole iterated scheme is equivalent to the Crank–Nicolsen method. We choose a global time step for the statistical model according to the expression

$$\Delta t = \min\left(\eta_{\rm Disc} \frac{X}{\dot{X}}, X \in \{\Sigma, T_r, T_z\}\right) \tag{291}$$

where the hybrid code (see next section) applies an additional discretisation in powers of two to achieve a better synchronisation with the N–body code component.

17 BRINGING THE TWO SCHEMES TOGETHER: THE HYBRID CODE

We introduced two different methods to solve the planetesimal growth problem. On the one hand, we modified NBODY6++, which has been used so far mainly for the simulation of stellar clusters, to adapt it to the special requirements of a long-term integration of planetesimal orbits. On the other hand, we developed a new statistical code with a consistent evolution of the velocity dispersion, the capability to treat spatial inhomogeneities and a thoroughly constructed collision treatment. Neither of the two approaches is powerful enough to provide a complete and accurate description of the planetesimal problem, since each method is confined to a certain range of the particle number. However, these restrictions are complementary in the sense that each method covers a regime where the other method fails. This intriguing relation stimulated the construction of a hybrid code which combines the benefits of both methods.

The basic idea is to introduce a transition mass $m_{\rm trans}$, which separates the two mass regimes. Particles with a lower mass are treated by the statistical model, whereas larger particles belong to the N–body model. Though both parts are clearly divided in different mass ranges, they are connected by various interdependencies:

(i) Direct collisions between particles lead to a mass exchange. One process is the accretion of small particles by N–body particles, but agglomeration within the statistical model can also produce particles larger than the transition mass. This requires the generation of new N–body particles. Energetic impacts may erode larger particles, so a corresponding particle removal is also required for consistency.

(ii) Mutual scatterings among N–body particles and smaller planetesimals transfer kinetic energy. While energy equipartition leads to a systematic heating of the smaller field planetesimals, a consistent treatment has to include both transfer directions.

(iii) Accretion and scattering by the N–body particles induce spatial inhomogeneities or even gaps in the planetesimal component, if the particles have grown massive enough. Likewise, the small particles could induce some structure in the distribution of the N–body particles. Since the spatial structure is dominated by the stirring from few protoplanets, we neglect the latter process.

Fig. 15 summarises this brief overview of the interactions between the two code components in a schematic diagram. The following sections explain the implementation of each interaction term in more detail.

17.1 Mass Transfer

An N–body particle accretes smaller particles in its vicinity. We already derived expressions which describe agglomeration within the statistical model, so it is manifest to apply these formulæ to derive the accretion rate of an N–body particle.

Most of the material is accreted within the cross-sectional area σ (see Eq. 267), but the finite eccentricity of an orbit extends the accessible radial feeding zone. Thus

we assign the following surface density to each particle

$$\Sigma(a) = \frac{M}{2\pi a\sqrt{2\pi}l} \exp\left(-\frac{(a-a_0)^2}{2l^2}\right) \tag{292}$$

$$l^2 = \sigma/\pi + \frac{1}{2}a^2e^2 + T_r/\Omega^2 \tag{293}$$

by smearing it out over its feeding zone. T_r is the radial velocity dispersion of particles in the statistical model with semimajor axis a. This density distribution is projected onto the radial grid to calculate the accretion rate. As the time step of the statistical model is much larger than the regular step of an N–body particle, the particle mass update is synchronised with the statistical integration. The projection technique allows the calculation of the accretion rates in a simple way, which gives the right size of the feeding zone and the proper total accretion rate.

Particle generation is included in the following way: A "virtual" mass bin is introduced as the boundary between the statistical grid (denoted by the dashed area in Fig. 15) and the N–body component. Its sole task is to store mass and kinetic energy that drives the statistical model towards higher masses. If the mass content exceeds one $m_{\rm trans}$, a new particle is created with inclination and eccentricity according to the stored velocity dispersions.

The masses of the N–body particles are regularly checked to detect any particle which dropped below the transition mass. While this procedure would remove the particle and transfer the associated quantities back to the grid, we never observed such a particle erosion.

17.2 Disc Excitation

The projection of an N–body particle onto the grid with the help of a proper weight function is also useful for the calculation of the disc excitation due to stirring by the larger particles. Since the Hill radius sets the proper length scale for planetesimal encounters, the weight function is modified to

$$\Sigma(a) = \frac{M}{2\pi a\sqrt{2\pi}l} \exp\left(-\frac{(a-a_0)^2}{2l^2}\right)$$

$$l^2 = R_{\rm Hill}^2 + \frac{1}{2}a^2e^2 + T_r/\Omega^2 \tag{294}$$

where T_r is the radial velocity dispersion of the heated planetesimal component. The velocity dispersion of the stirring N–body particle is (in accordance with Eq. 11 and Eq. 12):

$$T_{r,0} = \frac{1}{2}(\Omega a_0)^2 e^2 T_{z,0} = \frac{1}{2}(\Omega a_0)^2 i^2 \tag{295}$$

We employ the orbital elements as mediators between the fast varying instantaneous position and velocity of a particle and the slow evolution of the statistical model, which operates on a longer relaxation timescale. In virtue of the projection of the particle, we readily apply the standard interaction terms (see section 16) to evaluate the additional heating due to the presence of N–body particles.

17.3 Pseudo–Force

While an N–body particle is moving through the disc, it also interacts gravitationally with the particles in the statistical model. The collective effect of all these encounters leads to a change in the orbital elements of the N–body particle. Again, we project the N–body particle onto the grid and evaluate the stirring rates $\dot{T}_r$ and $\dot{T}_z$, which correspond to a change in the orbital elements:

$$\frac{d}{dt}e^2 = \frac{2\dot{T}_r}{(\Omega a_0)^2}$$

$$\frac{d}{dt}i^2 = \frac{2\dot{T}_z}{(\Omega a_0)^2} \tag{296}$$

These time derivatives of eccentricity and inclination are translated to a pseudo-force that effects the desired change of the orbital elements. We chose the ansatz

$$\mathbf{F}_{x,y} = C_r(\mathbf{v}_{x,y} - (\mathbf{v}_K)_{x,y})$$

$$\mathbf{F}_z = C_z\mathbf{v}_z \tag{297}$$

where $\mathbf{v}_K$ is the local Keplerian velocity. In addition, we tried a simpler expression

$$\mathbf{F}_{x,y} = 2C_r\mathbf{r}_{x,y}\frac{\mathbf{r}\cdot\mathbf{v}}{r^2}$$

$$\mathbf{F}_z = C_z\mathbf{v}_z \tag{298}$$

without any significant differences in the accuracy or the simulation outcome. The proper friction coefficients are:

$$C_r = \frac{\dot{T}_r}{2T_r}$$

$$C_z = \frac{\dot{T}_z}{2T_z} \tag{299}$$

Since the relevant quantities are the time derivatives of the orbital elements, any other pseudo-force is also applicable. Though this approach yields the right mean change of the orbital elements, it lacks the statistical fluctuations from the particle disc. Hence the distribution of the orbital elements of the N–body particles is artificially narrowed, which is especially important when the N–body particles and the statistical particles have a comparable mass. As the mass contrast between the two code parts is quite significant in planet formation simulations, it is safe to neglect the fluctuating part without major restrictions on the realism of the simulations.

The friction coefficients C_i are kept constant between two integration steps of the statistical model. While a more frequent update of the coefficients would be easily possible, a regular update on the basis of the statistical time step is accurate enough. Moreover, each update poses a considerable computational effort (roughly equivalent to 1000 force evaluations), so our approach also saves valuable computational time.

17.4 Spatial Structure

The first insight into planetesimal formation was obtained by the particle–in–a–box method, which invokes the underlying assumption that the planetesimal disc stays homogeneous throughout the protoplanet growth (Greenberg et al. 1978,see e.g.). While few large bodies introduce some coarse-graininess of the surface density, all smaller bodies are assumed to be evenly spread in the disc. Research on the interaction of protoplanets showed that this is an oversimplification, as bodies that are massive enough could open gaps in their vicinity (Lin & Papaloizou 1979; Rafikov 2001,see

No	Σ	Δa	N	$N_{\rm rad}$	e^2/h^2	i^2/h^2	m	Type
G1	1.1251×10^{-6}	0.2	1406	–	0.00135	0.00135	1×10^{-9}	N–body
	Perturber	—	1	–	$e = 6.1 \times 10^{-5}$	$i = 3.2 \times 10^{-5}$	1×10^{-7}	
G2	1.1251×10^{-6}	0.2	–	201	0.00135	0.00135	1×10^{-9}	Statistic
	Perturber	—	1	–	$e = 6.1 \times 10^{-5}$	$i = 3.2 \times 10^{-5}$	1×10^{-7}	

Table 10. Parameters of the statistical and the N–body gap simulation. The perturber is placed at the centre of the ring.

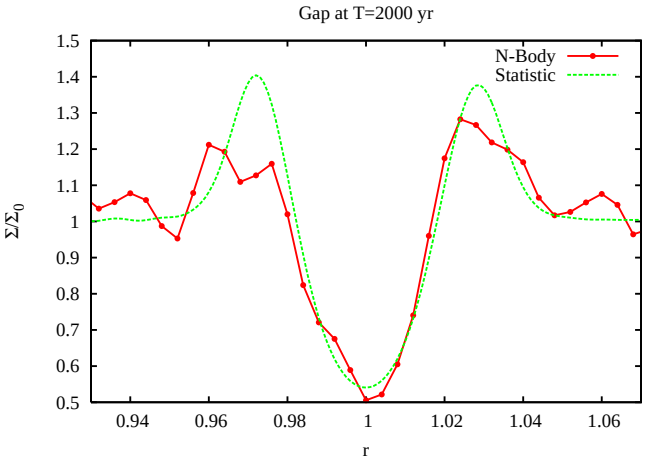

Figure 16. Gap opening in a planetesimal disc. The gap is fully developed after 2000 years. Table 10 summarises the initial conditions for the comparative runs.

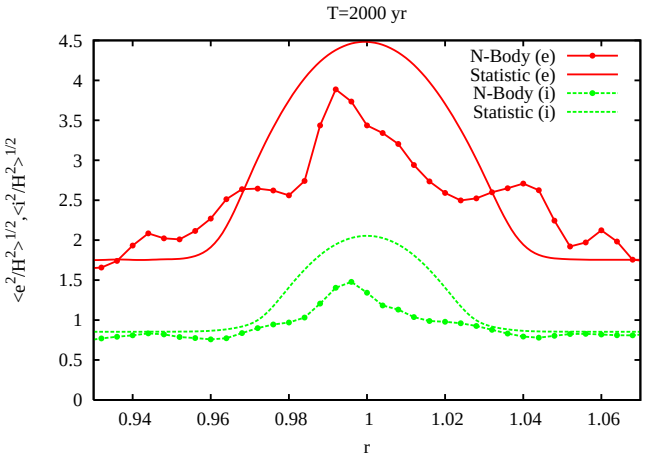

Figure 17. Mean square eccentricity and inclination of the smaller planetesimals in terms of the reduced Hill radius H of the protoplanet according to simulation G1 (N–body) and G2 (Statistic).

e.g.). Gap formation does not only change the overall surface density, but also controls the accretion onto the protoplanet through the amount of planetesimals in the feeding zone. If gap formation is too effective, the growth of the protoplanet may well stop before the isolation mass is reached. Hence any hybrid code should provide a framework that allows this mechanism to operate. A necessary condition is a radial density grid with a sufficient resolution to describe possibly emerging gaps. A too low resolution suppresses local perturbations from the protoplanets by a simple averaging, thus inhibiting the formation of any spatial inhomogeneities. A second requirement is that the interaction terms relating statistical model and N–body model include the local interaction between particles and the statistical component in a proper way.

Our hybrid approach includes gap formation implicitly through the diffusive terms. A protoplanet heats only the planetesimals in its vicinity (defined by the heating kernel), thus also increasing locally the diffusion coefficient. Hence the surface density drops due to outward diffusion of the planetesimals, given that the protoplanet is massive enough. The minimum gap opening mass is set by the condition that the protoplanet controls the random velocities of the field planetesimals in its heating zone (see Eq. 30), which is equivalent to the independently derived gap formation criterion (compare Eq. 32).

Although our algorithm invokes a simplified picture of the protoplanet–planetesimal interaction, it is surprisingly accurate with respect to the width of the forming gap and the opening criterion. Fig. 16 shows a simulation which examines the accuracy of our approach. The overall performance of the statistical code is quite remarkable, except a significant overestimation of the surface density at the gap boundary compared to the N–body model. This deviation is due to the improper treatment of strong planetesimal–protoplanet encounters, which exceed the diffusive approximation. Moreover, the higher concentration of planetesimals near the gap boundary leads to an additional overestimation of the velocity dispersion of the smaller planetesimals in the statistical calculation (see Fig. 17). While the comparison with the N–body calculation clearly indicates a necessary improvement of the treatment of spatial inhomogeneities, our approach catches the main features of gap formation.

17.5 Transition Mass

Since the inventory of the new hybrid code is now completed, we turn to the specification of the transition mass $m_{\rm trans}$. The mass boundary between statistical and N–body part has a major influence on the realism and the speed of the simulation. On the one hand, optimisation with respect to speed favours a large transition mass, whereas a reasonable resolution of the transition between the two components introduces some upper limit.

Hence we identify first the set of large masses, which controls the velocity dispersion of the disc, since these objects are also possible candidates for gap opening. The inspection of all involved stirring terms gives approximately

the inequality:

$$\int_0^{m_{\rm trans}} \frac{d\Sigma}{dm} m dm < \int_{m_{\rm trans}}^{\infty} \frac{d\Sigma}{dm} m dm \qquad (300)$$

While this is a necessary condition to select all potential major perturbers, criterion 300 does not imply that all particles in the selected mass range exert indeed a strong influence on the disc. The number of possible gaps – and therefore the number of perturbers associated with them – is ultimately limited by the available space. Thus we integrate the area of all potential gaps (width $\approx f_\Delta R_{\rm Hill}$) and normalise it to the total disc area:

$$f_C \approx \int_{m_{\rm trans}}^{\infty} f_\Delta \frac{d\Sigma}{dm} \frac{2\pi a R_{\rm Hill}}{m} dm \qquad (301)$$

If the covered fraction f_C is much larger than one, it is possible to increase the transition mass until the condition

$$f_C \lessapprox 1 \qquad (302)$$

is fulfilled. Of course, condition 300 and 302 defined only an upper limit of the transition mass, so the adaptation of a lower value is also possible. Though there are two reliable conditions at hand, the transition mass is still a function of time owing to the time evolution of the density $\Sigma(m)$. Therefore we chose a priori a fiducial value of the transition mass, run the simulation and conduct an a posteriori check, whether the initial choice matches our requirements at any evolutionary stage of the disc. A reliable value for a solar system analogue at 1 AU is

$$m_{\rm trans} \approx 3 \times 10^{-11} M_\odot \qquad (303)$$

which restricts the number of N–body particles to a tractable amount. Later stages would allow an even larger transition mass, but the current hybrid code does not include any dynamical adjustment of the transition mass at runtime.

17.6 Boundary Conditions

Any numerical simulation is limited to a finite simulation volume and a finite time interval. Therefore it is mandatory to introduce proper boundary conditions which provide a reasonable closure of the simulation volume.

While boundary conditions with respect to time are the familiar initial conditions, the choice of the spatial boundary conditions for the various involved quantities depends on the problem at hand and the type of the boundary. A simulation boundary can be due to physical reasons (like walls of a concert hall, surface of a terrestrial planet) or simply due to a limitation in computational power that inhibits the complete numerical coverage of the problem.

The current capability of the hybrid code sets limits on the radial range as well as on the covered mass range, which a simulation can handle in a reasonable time. Hence we have to introduce artificial boundaries in radius, and a lower limit for the mass grid.

Any migration process couples the evolution of a local ring area in the planetesimal disc to the evolution of the whole disc. Inward (or outward) migrating material also transports information on the radial zone where the material originated from. As this information is not available within the frame of a local simulation, any choice of the boundary condition alters the evolution to some extend.

However, we focus on a formation stage where migration is not a dominant process, but provides only removal of the smaller collisional fragments. Thus we apply closed boundary conditions for the outer and inner radius of the ring area (i. e. all fluxes vanish at the boundary), and an open boundary for the lower end of the mass range. While these conditions exclude the study of migrational processes, we gain clearer insight into the protoplanet growth.

18 DISCUSSION AND CONCLUSIONS

The formation of planetary systems represents a challenge from a numerical standpoint. The dynamical problem spans over many orders of magnitudes in length and demands the combination of different techniques. We have presented a composite algorithm that brings together the advantages of direct-summation tools and statistics for the description of the planetesimal disc. Direct-summation $N-$body techniques have been around for some decades and have proven their accuracy in a very large number of studies of stellar clusters such as galactic nuclei and globular and open clusters. We deem it to be the numerical tool to integrate the motion of the bodies for the very precise integration of the orbits and treatment of close encounters. Typically, in a simulation of a stellar system, the energy is conserved in each timestep by $E/\Delta E \sim 10^{-11}$ (where E is the total energy and ΔE the difference between the former and current total energy for a specific time), so that even if we integrate for a long time the cluster, the accumulated energy error is negligible. Nevertheless, porting the numerical tool to the problem of planetary dynamics is not straighforward and requires important modifications and additions. In this work we present them in detail: the neighbour radius selection for the protoplanets, the Hermite iteration and we introduce for the very first time the new extended Hermite scheme, since the usual Hermite scheme is not sufficient to integrate planetesimal orbits accurately enough. Then we bring in new forces to the problem, namely the introduction of the central potential of the star, as well as the drag forces, which depend on the gas density and size of the planetesimals. Hence, the regularisation scheme, crucial to exactly integrate the close encounters, has to be accordingly modified. We then introduce the disc geometry and discuss the required changes to the neighbour scheme and prediction, as well as the communication algorithm and block size distribution.

For the statistical description of the planetesimal disc we employ a Fokker-Planck approach. We include dynamical friction, high- and low-speed encounters, the role of distant encounters as well as gas and collisional damping and then generalise the model to inhomogenous discs. We then describe the combination of the two techniques to address the wole problem of planetesimal dynamics in a realistic way via a transition mass to integrate the evolution of the particles according to their masses.

In particular, we introduce and describe the extended Hermite scheme, which reduces the the energy error by three orders of magnitude with the same number of force evaluations, compared to the standard version of NBODY6++.

While the implementation and some code details are

No.	Σ	Δa	N	$N_{\rm rad}$	e^2/h^2	i^2/h^2	m	Type
T1a	1.1251×10^{-6}	0.02	1000	–	0.04	0.01	1.41×10^{-10}	N–body
T1b	1.1251×10^{-6}	0.02	500	10	0.04	0.01	1.41×10^{-10}	Hybrid
T1c	1.1251×10^{-6}	0.02	–	10	0.04	0.01	1.41×10^{-10}	Statistic
T2a	0.5626×10^{-6}	0.08	800	–	4	1	5×10^{-10}	N–body
	0.5626×10^{-6}		200	–	4	1	2×10^{-9}	
T2b	0.5626×10^{-6}	0.08	–	10	4	1	5×10^{-10}	Hybrid
	0.5626×10^{-6}		200	–	4	1	2×10^{-9}	
T3	Safronov	–	–	–	–	–	–	Statistic
T4a	1.1251×10^{-6}	0.02	10.000	–	4	1	1.41×10^{-11}	N–body
T4b	1.1251×10^{-6}	0.02	–	10	4	1	1.41×10^{-11}	Hybrid
T4c	1.1251×10^{-6}	0.02	–	10	4	1	1.41×10^{-11}	Statistic
T5	1.8789×10^{-6}	–	–	–	620	155	2.4×10^{-15}	Statistic

Table 11. Parameters of all test simulations. The transition mass in T4b is $m_{\rm trans} = 3.1 \times 10^{-10}$ Only simulations T3, T4a–T4c and T5 include collisions. All values are scaled to $M_c = G = r_0 = 1$.

newly introduced to the field of planet formation simulations, the first hybrid approach was developed in the early 90's. Spaute et al. (1991) (further improved in Weidenschilling et al. 1997) constructed a hybrid code with a statistical component to treat the smaller particles and a special treatment for the larger particles. A statistical model covers the field planetesimals with the help of a distribution function (similar to Wetherill (1989)), whereas the larger particles are individually stored and characterised by mass, semimajor axis, eccentricity and inclination. While the interaction between these single particles and the statistical component is expressed by standard viscous stirring and dynamical friction terms, perturbations among the single particles are equated in a different way. First, the probability of an encounter of two neighbouring particles is calculated. This probability is used in a second step to decide whether a (numerically integrated) two-body encounter of the neighbouring particles is carried out to derive the change in the orbital elements. Though these two well-defined code components justify to speak about a hybrid approach, the Monte–Carlo like integration of the largest particles is still closely related to a statistical treatment.

A modified N–body approach is used in the work of Levison & Morbidelli (2007). Their method covers the largest particles by a direct N–body code, which includes the smaller particles as "tracer" particles. The term "tracer" indicates that each particle represents a whole ensemble of planetesimals. In a similar line of approach and inspired by this idea, Levison et al. (2010) modified a symplectic algorithm, Symba, to study the formation of giant planet cores. However, they made some assumptions in order to calculate the gravitational interaction between the planetesimals. In particular, they ignored totally close encounters between planetesimals.

Ormel & Spaans (2008) present in their work a scheme based on Monte Carlo techniques to cover the vast range of sizes. For this, they assign more resolution to those particles that are more relevant to the interactions, typically the largest bodies. Smaller particles are grouped and treated collectively, which means that they all share the same mass and structural parameters. This classification is done in accordance to the "zoom factor", a free paramenter. Later, Ormel et al. (2010a) presented an detailed comparison of their Monte Carlo code with other techniques, in particular with pure direct-summation $N-$body results and other statistical studies and found that system leaves the runaway at a larger radius, in particular at the outer disc. With their simulations, the authors propose a new criterion for the runaway growth-oligarchy transition: from several hundreds of km in the inner disk regions up to a thousand km for the outer disc (Ormel et al. 2010b).

Bromley & Kenyon (2006) published a description of a hybrid method with a basic approach similar to our work. They employ two velocity dispersions and the surface density of the planetesimals to describe the planetesimal system. The statistical component includes migration of the planetesimals and dust particles due to gas drag and Pointing–Robertson drag. In contrast to our approach, they did not include mass transport due to the diffusion of the planetesimals, which precludes the study of spatial structures induced by the protoplanets. One must note also that their method uses the standard discretisation of the collisional flux (see Eq. 282) and thus restrict the spacing factor to $\delta \lesssim 1.25$ (Kenyon 1998). Bromley & Kenyon (2006) chose a set of test calculations which focused less on the technical aspects of their method, but on an overall comparison with a selected set of standard works on planet formation. Their test simulations are in good agreement with the references simulations, thus indicating a comparable quality of the method. Four years later, the authors presented an updated version of their code for planet formation. The new characteristics of the code included 1D evolution of the viscous disc, gas ac-

cretion on to massive cores, as well as accretion of small particles in planetary atmospheres (Bromley & Kenyon 2010).

While a variety of hybrid approaches emerged over the past years, this technique is still far from a routinely application and is still challenged by many open issues. Hybrid codes bear the potential to address the dynamical evolution of a whole planetary system, the later stages of protoplanet formation initiate a strong interaction with the gaseous disc, which may require more diligence than the inclusion of a few additional interaction terms. However, the development is picking up speed, which places our work in a good position for further research.

APPENDIX A: CENTRAL FORCE – DERIVATIVES

Central force $\mathbf{F}$ per mass (i. e. acceleration) and its time derivatives are:

$$\begin{aligned}
\mathbf{F} &= -\frac{\mathbf{x}M}{x^3} \\
\mathbf{F}^{(1)} &= -\frac{\mathbf{v}M}{x^3} - 3A\mathbf{F} \\
\mathbf{F}^{(2)} &= -\frac{\mathbf{a}M}{x^3} - 6A\dot{\mathbf{F}} - 3B\mathbf{F} \\
\mathbf{F}^{(3)} &= -\frac{\dot{\mathbf{a}}M}{x^3} - 9A\mathbf{F}^{(2)} - 9B\mathbf{F}^{(1)} - 3C\mathbf{F} \\
\mathbf{a} &= \dot{\mathbf{v}} \\
A &= \frac{\mathbf{x}\cdot\mathbf{v}}{x^2} \\
B &= \frac{v^2}{x^2} + \frac{\mathbf{x}\cdot\mathbf{a}}{x^2} + A^2 = \dot{A} + 3A^2 \\
C &= \frac{3\mathbf{v}\cdot\mathbf{a}}{x^2} + \frac{\mathbf{x}\cdot\dot{\mathbf{a}}}{x^2} + A(3B - 4A^2)
\end{aligned} \tag{A1}$$

The $\mathbf{F}^{(i)}$ denote the central force and its time derivatives, whereas $\mathbf{a}$ and $\dot{\mathbf{a}}$ refer to the total acceleration of the particle. The assumption that $\mathbf{x}, \mathbf{v}, \mathbf{a}$ and $\dot{\mathbf{a}}$ are independent of each other allows the derivation of averaged expressions for particle–particle interactions:

$$\begin{aligned}
\langle(\mathbf{F})^2\rangle &= m^2\frac{1}{x^4} \\
\langle(\mathbf{F}^{(1)})^2\rangle &= m^2\frac{2v^2}{x^6} \\
\langle(\mathbf{F}^{(2)})^2\rangle &= m^2\left(12\frac{v^4}{x^8} + 2\frac{a^2}{x^6}\right) \\
\langle(\mathbf{F}^{(3)})^2\rangle &= m^2\left(144\frac{v^6}{x^{10}} + 126\frac{a^2v^2}{x^8} + 2\frac{\dot{a}^2}{x^6}\right)
\end{aligned} \tag{A2}$$

We combine these expressions with Aarseth's time step formula to derive the regular time step as a function of the neighbour sphere radius R_s:

$$\begin{aligned}
\Delta t_{\rm reg} &\approx \sqrt{\eta_{reg}}\frac{R_s}{\bar{v}}\frac{1}{1+\sqrt{R_s/R_0}} \\
R_0 = 4\frac{\bar{v}^2}{a} &\approx \frac{4\bar{v}^2\bar{r}^2}{Gm} = 4\frac{\bar{r}^2}{r_{\rm close}}
\end{aligned} \tag{A3}$$

$\bar{r}$ is the average particle distance and $r_{\rm close}$ is the impact parameter for a 90–degree deflection.

APPENDIX B: SCALABLE COLLISIONS FLUX

The mass flux according to the perturbation equation 161 is:

$$\begin{aligned}
F_p = &-\iint (n(m_2)\Delta n(m_1) + n(m_1)\Delta n(m_2)) \\
&\sigma(m_1)v(m_2)m_1 f_m(m_1/m,\epsilon)dm_1dm_2 \\
&= F^{(1)} + F^{(2)}
\end{aligned} \tag{B1}$$

Firstly, we employ the substitution

$$\begin{aligned}
m_1 &= mx_1 \\
m_2 &= m_0\left(\frac{m_1}{m_0}\right)^{\frac{1+\alpha}{1+2w}}\left(\tilde{S}\right)^{\frac{1}{1+2w}}\epsilon^{\frac{1}{1+2w}}
\end{aligned} \tag{B2}$$

to solve for the partial flux $F^{(1)}$:

$$\begin{aligned}
F^{(1)} &= -n_0^2m_0^3\sigma_0v_0\int g(mx_1)F_1(x_1)dx_1 \\
F_1(x_1) &= \tilde{S}^{-k'}\int \epsilon^{-\frac{w+s+3+\alpha}{2+\alpha+2w}}\frac{f_m(x_1,\epsilon)}{x_1(1+2w)}d\epsilon
\end{aligned} \tag{B3}$$

The second contribution $F^{(2)}$ requires a slightly different transformation:

$$\begin{aligned}
m_1 &= mx_1\epsilon^{-1/(1+\alpha)} \\
m_2 &= m_0\left(\frac{mx_1}{m_0}\right)^{\frac{1+\alpha}{1+2w}}\left(\tilde{S}\right)^{\frac{1}{1+2w}}
\end{aligned} \tag{B4}$$

Thus the partial flux $F^{(2)}$ is:

$$\begin{aligned}
F^{(2)} &= -n_0^2m_0^3\sigma_0v_0\int g(m_2)F_2(x_1)dx_1 \\
F_2(x_1) &= \tilde{S}^{-k'}\int \epsilon^{-\frac{w+s+3+\alpha}{2+\alpha+2w}}\frac{f_m(x_1\epsilon^{-1/(1+\alpha)},\epsilon)}{x_1(1+2w)}d\epsilon
\end{aligned} \tag{B5}$$

We change to a new set of logarithmic coordinates

$$u = \ln(m/m_0) \qquad u_1 = \ln(x_1) \qquad \tilde{s} = \frac{\ln(\tilde{S})}{1+\alpha} \tag{B6}$$

which transforms the total flux F_m to a convolution integral:

$$\begin{aligned}
F_p = &-n_0^2m_0^3\sigma_0v_0\int [g(u+u_1)G_1(u_1) + \\
&g\left(p(u+u_1+s_1)\right)G_2(u_1)]du_1
\end{aligned} \tag{B7}$$

$$p = \frac{1+\alpha}{1+2w} \tag{B8}$$

$p = 1$ refers to the already derived solution for self-similar collisions. Hence we expand Eq. B7 at $p = 1$ and retain only the zeroth-order moment of the fragmentation kernel:

$$\begin{aligned}
F_p = &-n_0^2m_0^3\sigma_0v_0\Big[g(u)G_{1,0} + (g(u) + \\
&u(p-1)\frac{\partial g}{\partial u})G_{2,0}\Big]
\end{aligned} \tag{B9}$$

This expression is equivalent to

$$\begin{aligned}
F_p = &-n_0^2m_0^3\sigma_0v_0(g(u)G_{1,0} + \\
&[g(u) + (p-1)(g(u) - g(0))]G_{2,0})
\end{aligned} \tag{B10}$$

where higher derivatives of $g(u)$ are neglected. Hence we recover the same functional form of the perturbed mass flux F_p as for self-similar collisions:

$$\begin{aligned}
F_p &= -n_0^2m_0^3\sigma_0v_0g(u)\left(G_{1,0} + pG_{2,0}\right) + \text{const.} \\
&\propto \tilde{S}^{-k'}
\end{aligned} \tag{B11}$$

APPENDIX C: COAGULATION EQUATION

While the success of a general approximation of the coagulation equation depends heavily on the used coagulation kernel, we nevertheless provide a more general approach to embed section 10 in a broader context. The standard coagulation equation is:

$$\begin{aligned} 0 &= \frac{\partial}{\partial t} mn(t,m) + \frac{\partial}{\partial m} F_m(t,m) \\ F_m &= -\iint n(t,m_1)n(t,m_2)\sigma(m_1,m_2) \\ &\quad v_{\rm rel} M_{\rm red}(m,m_1,m_2) dm_1 dm_2 \end{aligned} \tag{C1}$$

In virtue of our experience drawn from the perturbation expansion, we transform the coagulation equation to logarithmic coordinates

$$u = \ln(m) \tag{C2}$$

and employ the size distribution $g(u)$ relative to the steady-state solution $n_{\rm eq}(m)$:

$$\begin{aligned} 0 &= \frac{\partial}{\partial t} g(u,t) n_{\rm eq}(u) e^{2u} + \frac{\partial}{\partial u} F_u(t,m) \\ F_u &= -\iint g(t,u_1) g(t,u_2) K(u,u_1,u_2) du_1 du_2 \end{aligned} \tag{C3}$$

$K(u,u_1,u_2)$ is the properly transformed new coagulation kernel. $g(u)$ is expanded under the integral to arrive at a moment expansion of the flux F_u:

$$\begin{aligned} F_u &= -K_{00}(u) g(u)^2 - (K_{10}(u) + K_{01}(u)) g(u) \frac{\partial g}{\partial u} + \dots \\ K_{ij} &= \iint K(u,u_1,u_2) u_1^i u_2^j du_1 du_2 \end{aligned} \tag{C4}$$

Retaining only the leading order terms, we recover an approximate coagulation equation which is similar to the inviscid Burgers' Equation [29]:

$$0 = \frac{\partial}{\partial t} g(u,t) n_{\rm eq}(u) e^{2u} - \frac{\partial}{\partial u} \left(K_{00}(u) g(u)^2 \right) \tag{C5}$$

ACKNOWLEDGMENTS

It is a pleasure to thank Sverre Aarseth, Cornelis Dullemond and Phil Armitage for comments on the manuscript. PAS thanks the National Astronomical Observatories of China, the Chinese Academy of Sciences and the Kavli Institute for Astronomy and Astrophysics in Beijing, for an extended visit, as well as the Aspen Center of Physics and the organizers of the summer meeting, where this work was finished. PAS expresses his utmost gratitude to Hong Qi, Wenhua Ju and Xian Chen for their hospitality during his stay in Beijing. He is also somehow marginally indebted with the 2011 winter strain of German H1N2, which allowed him to skip all possible duties and focus for an extended period of time on this work at home. RS acknowledges support by the Chinese Academy of Sciences Visiting Professorship for Senior International Scientists, Grant Number 2009S1-5 (The Silk Road Project). The special supercomputer Laohu at the High Performance Computing Center at National Astronomical Observatories, funded by Ministry of Finance under the grant ZDYZ2008-2, has been used. Simulations were also performed on the GRACE supercomputer (grants I/80 041-043 and I/84 678-680 of the Volkswagen Foundation and 823.219-439/30 and /36 of the Ministry of Science, Research and the Arts of Baden-urttemberg). Computing time on the IBM Jump Supercomputer at FZ Jülich is acknowledged.

[29] This notion goes back to Burgers (1948), but the equation was already introduced by Bateman (1915).

REFERENCES

Aarseth S., 1985, Academic Press Orlando, 378
—, 1999, PASP, 111, 1333
—, 2003, Cambridge University Press
Adachi I., Hayashi C., Nakazawa K., 1976, Progress of Theoretical Physics, 56, 1756
Ahmad A., Cohen L., 1973, Journal of Computational Physics, 12, 389
Armitage P. J., 2010, ArXiv e-prints
Astakhov S., Lee E., Farrelly D., 2005, MNRAS, 360, 401
Bagatin A., Cellino A., Davis D., Farinella P., Paolicchi P., 1994, Planet. Space. Sci., 42, 1079
Balbus. S., 2003, Ann. Rev. Astron. Astrophys., 41, 555
Balbus S., Hawley J., 1998, Reviews of Modern Physics, 70, 1
Bateman H., 1915, Monthly Weather Review, 43, 163
Beauge C., Aarseth S. J., 1990, MNRAS, 245, 30
Beckwith S., 1996, Nature, 383, 139
Benz W., 1999, Icarus, 142, 5
Binney J., 1977, MNRAS, 181, 735
—, 1994, Princeton University Press, Princeton, New Jersey
Birn J., 1973, Astronomy and Astrophysics, 24, 283
Blum J., Wurm G., 2000, Icarus, 143, 138
Bodenheimer P., 1986, Icarus, 67, 391
Bodenheimer P., Hubickyj O., Lissauer J., 2000, Icarus, 143, 2
Bromley B. C., Kenyon S. J., 2006, AJ, 131, 2737
—, 2010, ArXiv e-prints
—, 2011, ArXiv e-prints
Burgers J., 1948, Adv. Appl. Mech., 1, 171
Cameron A. G. W., 1973, Space Science Reviews, 15, 121
Cazenave A., Lago B., Dominh K., 1982, Icarus, 51, 133
Chambers J., 2006, Icarus, 180, 496
Chandrasekhar S., 1942, Astrophysical Journal, 97, 255
Connaughton C., Rajesh R., Zaboronski O., 2004, Physical Review E, 69, 061114
Dohnanyi J., 1969, Journal of Geophys. Research, 74, 2531
Duncan M. J., Levison H. F., Lee M. H., 1998, 116, 2067
Glaschke P., 2003, in ESA Special Publication, Vol. 539, Earths: DARWIN/TPF and the Search for Extrasolar Terrestrial Planets, M. Fridlund, T. Henning, & H. Lacoste, ed., pp. 425–428
Goldreich P., Lithwick Y., Sari R., 2004, Astrophysical Journal, 614, 497
Goldreich P., Ward W. R., 1973, ApJ, 183, 1051
Greenberg R., 1991, Icarus, 94, 98
Greenberg R., Hartmann W., Chapman C., Wacker J., 1978, Icarus, 35, 1
Greenzweig Y., 1992, Icarus, 100, 440
Hasegawa M., 1990, Astronomy and Astrophysics, 227, 619

Hayashi C., 1981, Progress of Theoretical Physics Supplement, 70, 35
Hénon M., 1986, Celestial Mechanics, 38, 67
Hernquist L., Hut P., Makino J., 1993, ApJ Lett., 402, L85+
Hill G., 1878, American J. Math., 1, 5
Housen K., 1990, Icarus, 84, 226
Ida S., 1990, Icarus, 88, 129
—, 1992, Icarus, 96, 107
—, 1993, Icarus, 106, 210
Ida S., Guillot T., Morbidelli A., 2008, ApJ, 686, 1292
Ida S., Kokubo E., Makino J., 1993, MNRAS, 263, 875
Inaba S., Barge P., Daniel E., Guillard H., 2005, Astronomy and Astrophysics, 431, 365
Inaba S., Tanaka H., Nakazawa K., Wetherill G., Kokubo E., 2001, Icarus, 149, 235
Johansen A., Klahr H., Henning T., 2006, ApJ, 636, 1121
Johansen A., Oishi J. S., Mac Low M.-M., Klahr H., Henning T., Youdin A., 2007, Nat, 448, 1022
Kary D., Lissauer J., Greenzweig Y., 1993, Icarus, 106, 288
Kempf S., Pfalzner S., Henning T. K., 1999, Icarus, 141, 388
Kenyon S., 1998, Astronomical Journal, 115, 2136
Kobayashi H., Tanaka H., Krivov A. V., Inaba S., 2010, Icarus, 209, 836
Kokubo E., 1995, Icarus, 114, 247
—, 1996, Icarus, 123, 180
—, 1997, Icarus, 131, 171
Kokubo E., Ida S., 2000, Icarus, 143, 15
—, 2002, ApJ, 581, 666
Kokubo E., Kominami J., Ida S., 2006, ApJ, 642, 1131
Kokubo E., Yoshinaga K., Makino J., 1998, MNRAS, 297, 1067
Kornet K., Stepinski T., Różyczka M., 2001, Astronomy and Astrophysics, 378, 180
Kustaanheimo P. E., Stiefel E. L., 1965, J. Reine Angew. Math.
Larson R., 1970, MNRAS, 147, 323
Levison H., Morbidelli A., 2007, Icarus, 189, 196
Levison H. F., Thommes E., Duncan M. J., 2010, 139, 1297
Lin D., Papaloizou J., 1979, MNRAS, 188, 191
Makino J., 1988, Astrophysical Journal Supplement, 68, 833
—, 1992, PASJ, 44, 141
Masset F. S., Papaloizou J. C. B., 2003, ApJ, 588, 494
Mikkola S., 1997, Celestial Mechanics and Dynamical Astronomy, 67, 145
Mikkola S., Aarseth S. J., 1998, New Astronomy, 3, 309
Moore A., Quillen A. C., 2010, ArXiv e-prints
Nakagawa Y., H., N., 1983, Icarus, 54, 361
O'Brien D., 2003, Icarus, 164, 334
O'dell C., Wen Z., Hu X., 1993, Astrophysical Journal, 410, 696
Ohtsuki K., Nakagawa Y., Nakazawa K., 1990, Icarus, 83, 205
Ohtsuki K., Stewart G., Ida S., 2002, Icarus, 155, 436
Ormel C. W., Dullemond C. P., Spaans M., 2010a, ApJ Lett., 714, L103
—, 2010b, Icarus, 210, 507
Ormel C. W., Spaans M., 2008, ApJ, 684, 1291
Paardekooper S.-J., Baruteau C., Kley W., 2011, MNRAS, 410, 293
Papaloizou J., 2006, Rep. Prog. Phys., 69, 119
Paszun D., Dominik C., 2009, A&A, 507, 1023
Petit J.-M., 1986, Icarus, 66, 536
Plummer H. C., 1911, MNRAS, 71, 460
Rafikov R., 2001, Astronomical Journal, 122, 2713
—, 2003a, Astronomical Journal, 125, 922
—, 2003b, Astronomical Journal, 125, 906
Reipurth B., Jewitt D., Keil K. E., 2007, Lunar and Planetary Information Bulletin, issue 110, p. 25 (May 2007), 110, 25
Safronov V., 1969, NASA-TTF
Skeel R., 1999, Applied Numerical Mathematics, 29, 3
Spaute D., Weidenschilling S., Davis D., Marzani F., 1991, Icarus, 92, 147
Spurzem R., 1999, The Journal of Computational and Applied Mathematics, Special Volume Computational Astrophysics, 109, 407
Stewart G., 2000, Icarus, 143, 28
Tanaka H., Inaba S., Nakazawa K., 1996, Icarus, 123, 450
Wada K., Tanaka H., Suyama T., Kimura H., Yamamoto T., 2009, ApJ, 702, 1490
Ward W., 1986, Icarus, 67, 164
Weidenschilling S., Spaute D., Davis D., Marzari F., Ohtsuki K., 1997, Icarus, 128, 429
Weidenschilling S. J., 1977, MNRAS, 180, 57
Wetherill G., 1989, Icarus, 77, 330
—, 1993, Icarus, 106, 190
Wisdom J., 1991, Astronomical Journal, 102, 1528